DCIEM CR 2001-003

# THE USE OF ELECTRONIC CHARTS AND RADAR IN MARINE NAVIGATION

by

Don C. Donderi

Human Factors North Inc
118 Baldwin Street
Toronto, Ontario
M5T 1L6

Project Manager
D. C. Donderi
514 398-6130

PWGSC Contract No. W7711-9-7561/A

On behalf of

DEPARTMENT OF NATIONAL DEFENCE

As represented by

Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue West
Toronto, Ontario, Canada
M3M 3B9

DCIEM Scientific Authority
Sharon McFadden
Phone: 416 635-2189
November 2000

© HER MAJESTY THE QUEEN IN RIGHT OF CANADA (2000)
©SA MAJESTE LA REINE EN DROIT DU CANADA (2000)

As represented by the Minister of National Defence

# REPRODUCTION QUALITY NOTICE

**This document is the best quality available. The copy furnished to DRDCIM contained pages that may have the following quality problems:**

**: Pages smaller or Larger than normal**

**: Pages with background colour or light coloured printing**

**: Pages with small type or poor printing; and or**

**: Pages with continuous tone material or colour photographs**

**Due to various output media available these conditions may or may not cause poor legibility in the hardcopy output you receive.**

☒ **If this block is checked, the copy furnished to DRDCIM contained pages with colour printing, that when reproduced in Black and White, may change detail of the original copy.**

## EXECUTIVE SUMMARY

The main goal was to learn whether there was a difference between two ways to use marine electronic chart displays: with or without a radar overlay. Six mariners carried out nine simulated navigation exercises which were set under low-visibility conditions in the St. Lawrence Seaway, Halifax Harbour, and Come-by-Chance. The exercises varied in the complexity of the chart display and in the number of other ships encountered. Each mariner ran each exercise under one of three conditions: radar-chart overlay only, separate radar display only, or free use of either the radar-chart overlay or the separate radar display. Objective performance was measured by how closely mariners kept to the prescribed course, how closely they approached known hazards to navigation, and how fast they piloted their simulated vessel. Mariners also completed self-ratings of performance and difficulty after each exercise.

Objective performance was almost identical under the choice, overlay or separate display conditions. Thus there is no reason to prefer one display condition over the others for navigation in pilotage waters. Mariners self-rated their own performance as higher under the choice condition. They rated the exercises with simple chart displays, which were harbour approaches or departures in open waters (Come-by-chance and the Halifax approaches) as producing better performance, less demand and less frustration than the exercises with complex charts that were set in Halifax Harbour and the St. Lawrence Seaway.

Based on earlier work carried out on the same simulator, navigation in piloting waters is more accurate with electronic charts than with conventional paper charts, and based on the new work reported here, the choice of electronic chart display mode: separate radar and chart, or radar overlay on the chart, can be left to mariners' discretion.

Earlier research supported by DCIEM demonstrated that subjective complexity ratings of computer-generated graphic displays could be predicted from the size of the file generated when the JPEG file compression algorithm is applied to the bitmap file representing the display. Observers in one new study rated the complexity of the chart, radar and overlay images studied earlier, and the results were compared to the previous findings relating subjective complexity to compressed file size. The earlier finding was confirmed. Then two more sets of bitmap images: web pages and nature scenes, were tested in the same way, and the relationship between JPEG compressed file size and complexity ratings was again confirmed. This confirmation and extension of the earlier findings increases the probability of successfully completing a research program to develop an image-based quantitative measure that will predict human performance on electronic displays.

102-00841/





## SOMMAIRE

L'objectif principal consistait à déterminer s'il existe une différence entre les deux façons d'utiliser les affichages de cartes électroniques maritimes : avec ou sans superposition d'image radar. Six navigateurs ont effectué neuf exercices de navigation simulée dans des conditions de faible visibilité sur la Voie maritime du Saint-Laurent, au port de Halifax et à Come-by-Chance. Les exercices variaient quant à la complexité des cartes affichées et au nombre des navires rencontrés. Chaque navigateur a exécuté les différents exercices dans l'une des trois conditions suivantes : seulement avec image radar et carte électronique superposées, seulement avec image radar et carte électronique distinctes et avec le libre choix de la superposition ou de l'affichage distinct. On a mesuré les performances objectives en déterminant dans quelle mesure les navigateurs maintenaient le cap prescrit, à quelle distance ils se rapprochaient des risques connus pour la navigation et à quelle vitesse ils pilotaient leur navire simulé. Les navigateurs ont également effectué l'auto-évaluation des performances et des difficultés après chaque exercice.

Les performances objectives étaient presque identiques dans les conditions de libre choix, de superposition ou d'affichage distinct. Il n'y a donc aucune raison de privilégier une condition d'affichage pour traverser des eaux navigables. Les navigateurs ont jugé que leur propre rendement était supérieur dans la condition du libre choix. Ils ont souligné que les exercices faisant appel à des affichages de cartes simples, utilisées à l'entrée et à la sortie des ports en eau libre (Come-by-chance et Halifax), permettaient d'obtenir de meilleures performances et occasionnaient moins de contraintes et de frustrations que les exercices faisant appel à des cartes complexes, utilisées dans le port de Halifax et sur la Voie maritime du Saint-Laurent. Selon des travaux antérieurs exécutés avec le même simulateur, la traversée des eaux navigables est plus précise avec des cartes électroniques qu'avec des cartes ordinaires sur papier et, d'après les nouveaux travaux dont il est question ici, le choix du mode d'affichage des cartes électroniques, avec affichage distinct des images radar et des cartes ou avec superposition des images radar et des cartes, peut être laissé à la discrétion du navigateur.

Des recherches antérieures soutenues par l'IMCME ont démontré qu'il était possible de prévoir la complexité subjective des affichages graphiques générés par ordinateur en se fondant sur la taille du fichier produit lorsque l'algorithme de compression de fichier JPEG est appliqué au fichier en mode point représentant l'affichage. Dans le cadre d'une nouvelle étude, des observateurs ont évalué la complexité des cartes, images radar et images superposées examinées précédemment, et les résultats ont été comparés aux conclusions antérieures qui établissaient un lien entre le complexité subjective et la taille des fichiers comprimés. Les conclusions antérieures ont été confirmées. Par la suite, deux autres ensembles d'images en mode point, soit des pages Web et des scènes de la nature, ont été soumis à des essais semblables, et la relation entre la taille des fichiers comprimés JPEG et l'évaluation de la complexité s'est de nouveau confirmée. Cette confirmation et l'extension des conclusions antérieures augmentent la probabilité du succès d'un programme de recherche visant à mettre au point une mesure quantitative fondée sur les images, qui permette de prévoir le rendement humain obtenu avec des affichages électroniques.

## ETHICAL CERTIFICATION

This study, approved by the McGill University Research Ethics Board II, was conducted in conformity with the *Guidelines on Research Involving Human Subjects*, published by the Medical Research Council of Canada

## ACKNOWLEDGEMENTS

I thank the staff of the Centre for Marine Simulation, St. John's, Nfld, for their wholehearted cooperation and support, without which this research could not have been accomplished. I thank in particular Capt. Douglas Skinner, Interim Director of the Centre at the time this work was done, Capts. Bob Mercer and Fred Anstey, who chose the scenarios, scheduled the observers and collected the experimental data, and Jason Poole and James Green, Marine Institute students, who acted as data recorders and helmsmen during the study. Finally I thank Dr. Jacek S. Pawlowski, Research Director at the Institute, for his encouragement and cooperation in all aspects of the project. It was a pleasure working with all of them.

I thank Patricia Ann Turner for carrying out the two subjective rating experiments that comprise the second study reported here. She was responsible for all aspects of both experiments, including choosing the sample photographs and capturing the web images, printing and preparing the images for presentation, writing the experimental protocols, recruiting and training observers, collecting the data, and tabulating and filing the results. Her part of the work was performed flawlessly, and on schedule.

I thank Sharon McFadden, the DCIEM Scientific Authority, for her essential technical assistance at a vital early stage of this research, for the knowledge and experience that enabled her to contribute to explaining the present results, and for her strong encouragement to complete and to report these studies.

## TABLE OF CONTENTS

Executive Summary i

Ethical Certification ii

Acknowledgements ii

General Introduction 1
- Display Integration in Marine Navigation 1
- The Regulatory Environment for Electronic Charts 3
- What is Known about Electronic Chart Navigation 4
- Scope of the Present Work 6

Study 1: Piloting with Electronic Chart Displays 7
- Method 7
  - Simulator 7
  - Exercises 9
  - Display conditions 10
  - Mariners 10
  - Experimental design 11
  - Performance Measures 11
    - Objective measures 11
    - Mariner activity 12
    - Self-ratings 13
    - Written comments 13
- Results 14
  - Display conditions 14
  - Exercise types 15
    - Objective measures 15
    - Self-ratings 15
    - Activity 15
  - Mariners 22
    - Objective measures 22
    - Self-ratings 22
    - Activity 22
    - Summary of written comments 22
- Discussion 23

P515400.PDF [Page: 9 of 67]



## TABLE OF CONTENTS (COTD.)


Study 2: The Complexity of Visual Displays 25
General Introduction 25
Experiment 1: Replication of the Complexity Study 25
Method 25
Results 27
Experiment 2: Extension to New Image Sets 29
Method 29
Images 29
Observers 29
Design 29
Results 34
Compression algorithms 34
JPEG File size and complexity 34
Discussion 34

General Discussion and Conclusions 35
Reducing Overlay Display Clutter 36
Separate Language Layers 36
Independent Gain Control for Land Radar Returns 36
Future Developments in Electronic Navigation Displays 37
Flat-Panel Displays 37
Display Information Layers 38
Display Complexity Measurement 38

References 40

Appendix A. Description of Each Exercise 42

Appendix B. Mariner's Briefing Sheet for the Simulation Experiment 70

Appendix C. Mariner's Self-Rating Scale 73

Appendix D. Mariner's Written Comments 75

Appendix E. Letter Concerning Bridge Pier and ECPINS Chart 79

Appendix F. Complexity Rating Instruction and Data Form 81

Appendix G. Dissimilarity Derived from Complexity 87

## LIST OF TABLES

Table 1. Navigation Performance Comparison Between Paper Chart and Electronic Charts With or Without Radar Overlay: From Mercer and Honig (1994) 5

Table 2. Summary of the Six Exercises 9

Table 3. Overview of the Six Mariner Participants 10

Table 4. Performance Measures for Display Conditions and Exercise Types 16

Table 5. Individual Mariners' Performance 21

## LIST OF FIGURES

Figure 1. Plan of simulator wheelhouse 7

Figure 2. Photograph of the starboard (right-hand) corner of the wheelhouse 8

Figure 3. Photograph of an ECPINS display 8

Figure 4. Simulator control room 13

Figure 5a. Average cross-track deviation for display conditions and exercise types 17

Figure 5b. Self-rated performance for display conditions and exercise types 17

Figure 6. Objective performance measures for exercise types 18

Figure 7. Self-rating and activity measures for exercise types 19

Figure 8. Significant differences in navigator's performance on objective measures, activity measures and self-ratings 20

Figure 9. Relation between JPEG compressed file size and complexity from 1999 study and from the present replication 28

Figure 10. Four of the twenty nature photographs used in the replication study 30

Figure 11a. Two of the twenty web sites used in the replication study 31

Figure 11b. Two of the twenty web sites used in the replication study 32

Figure 12. Relationship between JPEG compressed file size and rated complexity for the website and natural scene images 33

## GENERAL INTRODUCTION

The past century has seen radical changes in the tools of marine navigation. The magnetic compass was followed by the gyrocompass, sextant and chronometer by Decca and Loran, and binoculars by radio and then by radar. And now, with differentially corrected global positioning satellites, it is possible to accurately locate a ship on an electronic chart in real time, to display navigation and piloting information on or alongside the chart, and to superimpose an accurately aligned radar image over the chart. Electronic chart systems, introduced to commercial marine navigation about ten years ago, are well on their way to completing the technical transformation of an ancient art.

These changes bring new opportunities and raise new questions for the mariner. Disregarding for the moment the vital questions of the availability and reliability of electronic chart systems, other important questions are: is electronic chart navigation objectively easier or safer than paper chart navigation? And since there are so many different ways of using electronic charts, which way is best? Since there are so many different possibilities and there are so many mariners, each mariner may eventually discover his or her own best choice among the many electronic combinations. Perhaps so; but when everyone learns by trial-and-error, no one learns from common experience. Applied research in electronic chart navigation is one technique for discovering and codifying common experience, perhaps more quickly and reliably than can be done through word-of-mouth and commercial competition.

Previous work carried out at the Centre for Marine Simulation in St. John's, Newfoundland (CMS) was intended to evaluate the performance differences between electronic chart and paper chart navigation in pilotage waters. The new work reported here was intended to study the differences between two methods of using electronic charts: with or without a radar overlay.

A second goal of the present work was to replicate and extend the finding from earlier studies that the subjective complexity of electronic navigation displays can be predicted from a simple quantitative measure of non-redundant information in the display.

### DISPLAY INTEGRATION IN MARINE NAVIGATION

A marine chart and a marine radar display are both scaled two-dimensional representations of what is on or near the surface of the water. Each type of display presents information that is redundant with and complementary to the information presented by the other. Radar and chart both show the presence of large, immoveable, above-surface landmarks or hazards to navigation like prominent buildings, bridges and piers. Charts show the planned location of buoys and aids to navigation while radar can often show the actual location of the same buoys and aids. Charts show the underwater location of invisible hazards like shallow water, rocks, reefs, wrecks and sub-sea cables, but cannot show recent or temporary hazards that developed after the last chart revision. Radar shows many ships and other transient hazards to navigation, but may miss small or

transient hazards that are not good radar reflectors. Neither chart nor radar can substitute for vision when visibility permits observation of the seaway, although when vision is limited, navigation proceeds cautiously, using charts and radar as guides.

The conventional paper chart (and an accompanying pilot book) is the traditional compendium of navigation information. Because the digital grain of a paper chart is fine, small print on a chart is readable, a range of colors and shades can be used to convey information about depth contours, land features, and navigation aids, and symbols and numbers convey information about depths, anchorages, buoys, ranges, landmarks and navigation lights, as well as compass variation and declination. On a paper chart, all of this information is conveyed simultaneously on the same surface; usually a large one up to a meter and a half in length and almost a meter in width. The chart table is the traditional home of the paper navigation chart as well as the navigator, who plots courses and bearings on the paper chart with pencil, straightedge and dividers. Marine surveys have been carried out with increasing accuracy and comprehensiveness since the seventeenth century, and until recently almost all of the information so gathered has been presented and used on paper charts. Marine navigation charts have a long history.

The introduction of marine radar in mid-century transformed plotting a ships' course with respect to fixed and moving objects, the essence of the art of pilotage, from a wholly visual task to an at least partly electronic one. Ranges and bearings from landmarks and moving targets could be calculated using radar and then transferred to the paper chart, or even plotted on the radar scope itself. Radar is a great help at night or in fog, but because radar does not display all potentially hazardous targets, and because radar navigation demands the same disciplined plotting of ranges and bearings as does visual navigation, radar does not immunize ships from collisions or groundings. It is as possible to make mistakes using radar as it is with visual navigation (Habberly, Shaddick & Taylor, 1986).

Commercial long-range radio positioning aids to replace sextant and chronometer also appeared in the decades following the Second World War. A numerical readout on a Loran or Decca receiver is associated with a predetermined position on a navigation chart. The information is either conveyed by chart lines that correspond to the numerical output of the system, or it is simply transformed to latitude and longitude coordinates and then plotted directly on a conventional navigation chart.

Mechanisms to integrate navigation information from multiple position sensors, as a way to replace the laborious and potentially inaccurate hand plotting of bearings and ranges on paper charts, have been under development since the nineteen-sixties. Automatic Loran pen position plotters were one such mechanism; another was an integrated plotting table that displayed a calculated ship's position on a paper chart as a projected spot of light (Istace, 1977). The marine navigation system designer's goal of integrating spatial information from multiple sensors is consistent with the generally high efficiency of the human visual system. Humans controlling the trajectory of moving objects, including ships, are engaging in visual-motor behavior. When visual-motor tasks are highly practiced, they become an efficient transaction between eye and body. This makes it easier to pay attention to other current problems that require listening and

talking to resolve (Wickens, 1992, pps 375-381). Therefore an efficient marine navigation system, whether based on paper or electronics, makes the navigator better able to deal with other aspects of managing a ship. The simplest imaginable marine sensor integration, as developed long before the advent of electronic charts (Millar and Hansford, 1983) places a constantly updated ship position symbol on a chart without the necessity of a repeated and laborious reference to ranges and bearings, dead reckonings, or Loran and GPS readouts. Automatic position updating supersedes the navigator's time-consuming paper-and-pencil task of just a generation ago, and increases the efficiency of this particular visual-motor control system.

However, current electronic chart systems do have important limitations. Even the best display screens have resolutions far below that of a paper chart, so the amount of information that can be read per degree of visual angle on an electronic chart is less than on paper charts. This forces the electronic chart designer to provide information in overlays, with a basic level of essential information required to be present at all times (IMO, undated). The use of overlays means that the chart has to be supplemented by a menu, and the user must either bear in mind (or be clearly reminded) about what type of information is on display, and what other types of information are available by changing the display mode. This imposes a memory burden that is not imposed by a paper chart, and may lead to information being ignored.

Electronic chart display screens are still relatively small, about 30 $cm^2$, nowhere near the more than 1 $m^2$ of a paper chart. This means that the total information presented on an electronic chart, as well as its density, is much less than on a paper chart. This forces the electronic chart user to either move the small chart aperture as the ship moves, maintaining a small ship-centered display, or to move the ship slowly over the stationary chart and then frequently jump the chart as the ship approaches the limit of the display.

A novel feature of the electronic chart system is its ability to superimpose a matched radar display over the chart. The radar is presented so that the chart information is visible simultaneously with, and spatially correlated with, the radar returns. The same scale and the same display mode (north-up or course-up) are imposed on both the chart and the radar display. The single topographical display eliminates the necessity (but not the choice) of viewing a separate radar display. Another useful feature of the display superimposition is that discrepancies between the redundant radar superimposition and the chart display alert the navigator to the fact that one or the other display is misaligned; in the event of a GPS positioning failure, the chart can be "map-matched" to the radar display to maintain it in alignment.

## The Regulatory Environment for Electronic Charts

International marine navigation standards as codified by IMO resolution A.817 (19) now recognize that certain compliant electronic chart displays, with provision for non-electronic backup, may satisfy the SOLAS treaty requirement that ships navigate with up-to-date charts. However, so long as a ship stocks the traditional supply of standard paper

navigational charts, any electronic navigation system may be used on its bridge, whether that system complies with the IMO resolution or not. During this period of technological transition, "let a thousand flowers bloom" is the order of the day, and many of these electronic chart systems will probably disappear in the face of competition.

## What is Known About Electronic Chart Navigation

The CMS carried out a simulation study in which conventional radar and paper chart navigation was pitted against an electronic chart display and a separate radar display, and against an electronic chart display combined with a radar overlay (Mercer and Hong, 1994). Six master mariners were first trained to use an electronic chart display (none had previous electronic chart experience). Then each of them was tested for his ability to navigate the same pre-plotted course through Halifax Harbour. The three conditions (paper chart, electronic chart, electronic chart with radar overlay) were each run under both good and poor visibility conditions, and the traffic pattern was varied randomly over each run. Objective performance measures included average cross-track deviation, closest point of approach to navigational hazards, and the number of rudder orders. The participants also self-rated their performance, their estimate of the demands imposed by each task, and their level of frustration in completing the task.

The two electronic chart conditions were different from the conventional chart navigation condition on every measure. The electronic chart displays produced lower cross-track deviations, closer to optimum distance between the own ship and oncoming ships, lower percentages of time spent monitoring the radar display, lower assessments of navigational workload, and lower levels of self-assessed task demand and frustration. These results were consistent under both the good and the poor visibility condition, but self-assessed workload and difficulty was greater under the poor-visibility condition. The results of this 1994 study are summarized in Table 1.

The information density limit on electronic charts can be overcome by presenting information in separate display layers that can be added or subtracted under user control. Kaufmann (1990) studied the differences in speed and accuracy of response to navigation questions that could be answered by reference to either a simplified (minimum information layer) or a complex (all information layers) static electronic chart display, and compared performance on both of the electronic chart displays to the responses made to the same topographic area presented as a complex paper chart. He tested nine observers (all cartographers or navigators) on one geographical area under all three conditions, and found that responses took less time, and there were fewer errors made, to the simple electronic chart display than to either the complex display or the paper chart. The differences among his groups did not reach conventional levels of statistical significance, but this was probably due to the small size of his observer group and the fact that they were tested on only one sample of each type of display.

Table 1. Navigation Performance Comparison Between Paper Chart and Electronic Charts With or Without Radar Overlay: from Mercer and Hong (1994).

| Measure | Exercise type | | |
|---|---|---|---|
| | Paper chart | ECDIS | ECDIS with overlay |
| | Objective measures | | |
| CTX | 2.2096 | -0.3382 | 0.0995 |
| Helm orders | 8.0833 | 13.0833 | 10.0833 |
| CPA | 0.0285 | -0.0059 | -0.0007 |
| Percent radar monitor | 35.89 | 15.90 | 15.30 |
| | Self-Ratings (TLX) | | |
| Navigation workload | 63.50 | 51.33 | 43.16 |
| | Good visibility | | |
| Mental | 2.66 | 0.50 | 0.83 |
| Physical | 3.16 | 0.50 | 1.00 |
| Time | 3.33 | 0.50 | 0.66 |
| Performance | 1.00 | 4.00 | 4.00 |
| Effort | 3.16 | 0.83 | 0.50 |
| Frustration | 2.66 | 0.33 | 1.50 |
| | Poor visibility | | |
| Mental | 4.66 | 2.83 | 3.50 |
| Physical | 4.83 | 2.66 | 2.82 |
| Time | 4.83 | 2.83 | 2.83 |
| Performance | 0.16 | 3.00 | 2.83 |
| Effort | 4.83 | 3.00 | 2.86 |
| Frustration | 4.33 | 2.66 | 3.50 |

CTX = Cross-track deviation, .1 nm, CPA = Closest Point of Approach, .1 nm.

Performance on static electronic charts with and without radar overlays was studied by Donderi (1999). Using a computer-based response technique similar to that employed by Kaufmann, college students and licensed mariners evaluated navigation statements, as rapidly and accurately as possible, by looking at either a separate pair of adjacent radar and chart displays or at a single overlay radar and chart display. The statements to be evaluated, e.g. "There is another ship to the left (port) of 'own ship.' ", were presented before the displays were seen. After the observer read the statement, he or she turned on the display and then responded as soon as possible by keying either "T" for true or "F" for false, to evaluate each statement. Fifty-two displays were drawn from electronic charts and radar simulations of four harbours. Accuracy was equal under the separate and overlay presentation conditions, but both the mariners and the college students spent less time looking at the overlay display, so efficiency (accuracy/time) was higher on the overlay displays.

Donderi also carried out a separate experiment to measure the subjective complexity of each display. College students rated the chart, radar and overlay displays for complexity, and then rated each chart and radar pair for dissimilarity. The overlay complexity measures were a linear function of chart complexity, radar complexity and chart-radar pair dissimilarity. When performance efficiency on separate and overlay displays was plotted against chart-radar dissimilarity, optimum efficiency was found at an intermediate level of dissimilarity.

Tullis (1983) reviewed research on the formatting of alphanumeric displays, and he also developed a quantitative measure of the structural information content of an alphanumeric display. In 1999, Donderi reported that there was no quantitative measure that corresponded to the subjective complexity of a graphic display like the electronic images used in his experiments. Subsequently, however, he found that the JPEG compression file size of a bitmap image file was highly correlated with the rated complexity of the bitmap image (Donderi & McFadden, 1999), thus introducing the possibility that the performance efficiency of an electronic chart display could be predicted from measurable physical characteristics of the display.

## Scope of the Present Work

The work reported here was planned to extend the knowledge gained through the work reported by Donderi (1999). The first goal was to compare performance on overlay or separate electronic chart displays under dynamic, as opposed to static, presentation conditions. The second goal was to replicate the subjective complexity measurements of the static chart, radar and overlay displays used in the previous research, as well as to replicate the discovery that subjective display complexity was a function of JPEG file compression size. The two goals were met through separate studies, the first carried out at the CMS, and the second at McGill University.

## STUDY 1: PILOTING WITH ELECTRONIC CHART DISPLAYS

### METHOD

#### Simulator

The marine simulator at the CMS consists of a ship's wheelhouse installed on a six degrees-of-freedom motion platform, surrounded by a 360° visual projection screen, and operated by a Norcontrol computer system. For this study, the wheelhouse was equipped with a Norcontrol DB 2000 ARPA radar console and an Offshore Systems Limited ECPINS v.1.4 electronic chart display system.

The ECPINS display system presents an electronic chart based on ENC data, oriented to and coordinated with the ships' position as displayed by an "own ship" mark on the electronic chart. The chart orientation and scale can be varied. The system allows the user to superimpose a transparent radar overlay at the same position, scale and orientation as the chart. The radar display remains oriented to the chart display as the ship moves.

The electronic chart and radar displays were located forward on the starboard (right-hand) side of the wheelhouse. A plan view of the simulator wheelhouse is shown in Figure 1. A general view of the wheelhouse as configured for this study is shown in Figure 2. The ECPINS display shown in Figure 3 was positioned in the location earlier occupied by a chart table (No. 4 in Figure 1). The radar display shown in Figure 3 was positioned at location No. 3 in Figure 1. During the exercise, the navigating officer carrying out the exercise issued engine and rudder commands to a helmsman (an experimental assistant, a student at the Marine Institute, St. John's) who carried them out on the steering stand, which for the duration of this study also contained the engine controls.

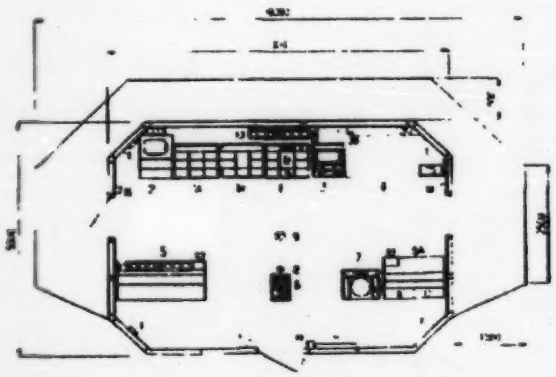


Figure 1. Plan of simulator wheelhouse layout. The ECPINS system replaced the chart table identified as No. 4.

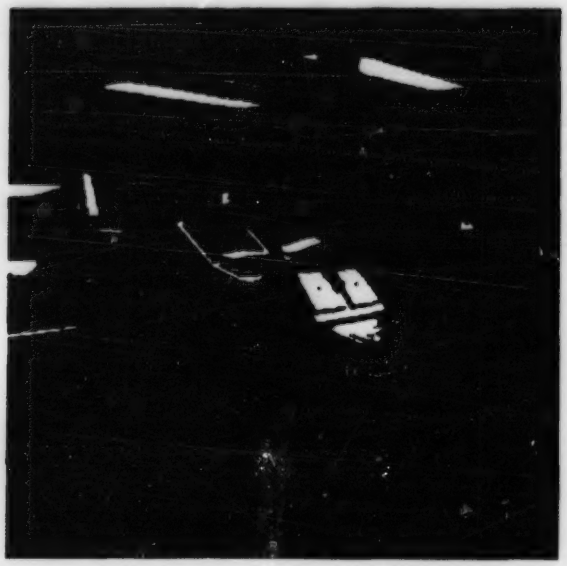


Figure 2. Photograph of the starboard (right-hand) corner of the simulator wheelhouse as configured for the study reported here, showing the ECPINS console at the front of the wheelhouse furthest to the right, the Norcontrol A2000 radar console to the left of it, the steering stand in the foreground, and an auxiliary radar display (not used by the mariner participants) located behind the ECPINS and Norcontrol displays.

Figure 3. Photograph of an ECPINS display, showing navigation and display information on the left, and an electronic chart on the right.

## Exercises

Nine navigation exercises were chosen by the staff of the CMS. The average exercise length as completed by the mariners was 42 minutes, and the range was from 36 to 52 minutes. The exercises were selected from the stored geographical database available to the CMS, and were chosen to meet two separate requirements. The first requirement was variation in chart and navigational complexity. Six exercises were chosen to take place in confined waters with complicated charts, and three exercises were chosen to take place in open waters with simple charts. The second requirement was variation in traffic density. Six exercises were carried out with between five and nine other moving ships visible or tracked on radar, and three exercises were carried out with only two or three other moving ships visible or tracked on radar. The nine exercises were grouped into three groups of three each. Group RC contained the three exercises with high traffic density and high chart complexity, group Rc contained the three exercises with high traffic density and low chart complexity, and group rC contained the three exercises with low traffic density and high chart complexity.

Each exercise included a pre-plotted course which was projected on the ECPINS display from the start position of the own ship symbol. The goal of the exercise was to follow the pre-plotted course as closely as possible, with due regard to the hazards to navigation, including oncoming and passing traffic. The navigating officer was free to vary the course and speed of the vessel at any time. In order to maximize reliance on the electronic displays, the visibility on all exercises was reduced to 0.4 nm by simulated fog, and the exercise was run without simulated communication to other ships or shore stations. Table 2 summarizes characteristics of each exercise. A more complete description of each exercise is found in Appendix A.

Table 2. Summary of the Nine Exercises

| Exercise type | Exercise no. | Moving targets | Exercise location and direction of ship movement |
|---|---|---|---|
| RC | 1 | 9 | Approaches to Halifax Harbour, inbound |
| | 4 | 4 | Halifax Harbour outbound from Bedford Basin |
| | 7 | 5 | St. Lawrence from Tousaint Island, up bound |
| Rc | 3 | 6 | Come-by-Chance outbound using traffic lane |
| | 6 | 5 | Inbound to Come-by-Chance |
| | 9 | 7 | Halifax outbound from south of Mars Rock buoy |
| rC | 2 | 3 | Halifax Harbour inbound to Bedford Basin |
| | 5 | 2 | St. Lawrence above Brockville Narrows, down bound |
| | 8 | 2 | St. Lawrence below Brockville Narrows, up bound |

## Display Conditions

The three display conditions were separate, overlay and choice. These terms refer to the way in which the radar display was used in association with the electronic chart display. In the separate condition, the radar display and the electronic chart display were always presented on separate but adjacent display consoles. The console locations are shown in Figures 1 (locations 3 and 4) and Figure 2 (the Norcontrol and ECPINS consoles). The radar superimposition feature of the electronic chart display was disabled. In the overlay condition, the navigator did not use the Norcontrol radar console at all. The radar information was always displayed as a transparent overlay on the electronic chart, and both the radar and the chart were viewed on the ECPINS console (location 4 in Figure 1; the far right-hand display in Figure 2). In the choice condition, the navigator was free to use either the separate radar display, or the transparent radar overlay on the electronic chart. He could turn the transparent overlay off or on at will, and could move between the radar display and the ECPINS display at will.

## Mariners

Six licensed mariners, all men, served as experimental observers. Their experience as licensed mariners ranged from eleven to 40 years. Four of them had worked at sea during the year preceding the study. All were associated with the CMS as instructors. Only one of them had previous practical experience with electronic chart systems, and he had not used a radar overlay. A brief overview of the mariners' qualifications is presented in Table 3.

Table 3. Overview of the Six Mariner Participants.

| Mariner No. | License | Years experience | Electronic chart experience | Sailed in last 12 months? |
|---|---|---|---|---|
| 1 | Master mariner | 20 | No | No |
| 2 | Master mariner | 19 | No | 4 months, Arctic trade |
| 3 | Master mariner, Tanker 1 & 2 | 40 | No | 2 months, Coastal trade |
| 4 | Master mariner, Tanker 2 | 14 | Two years, no radar overlay | 6 months, Research vessel foreign-going |
| 5 | ON I, Tanker 1 | 11 | No | No |
| 6 | ON I | 22 | No | 6 months, Offshore supply |

Each mariner was given a preliminary training session consisting of a tour of the wheelhouse and its facilities as arranged for this experiment, plus a familiarization exercise of approximately 45 minutes' duration that preceded the experimental exercises. The training session consisted of an exercise (entrance to St. John's Harbour) that was not used during the experiment. During the training exercise the navigator used the electronic chart equipment in all three configurations; separate display, overlay display, and choice of displays. Mariners were briefed in advance about the purpose of the study (Appendix B), and were encouraged to contribute their own written comments on the self-rating data sheets (Appendix C) that they completed following each run.

## Experimental Design

Following the training session, each mariner was tested on nine experimental exercises. There were three exercises of each type (RC, Rc and rC, see Table 2). One exercise of each type was randomly assigned to the choice display condition, one to the separate condition, and one to the overlay condition, for each mariner. Each mariner completed the exercises in a separately randomized order.

Each exercise was completed within a two-hour time period, which included a preliminary briefing on the exercise condition to be run (choice, overlay or separate), the exercise itself, and the post-exercise debriefing and self-rating. Mariners were never scheduled for more than two sessions per day; one in the morning and one in the afternoon. The study was completed between March 11 and March 24, 2000. Four exercises had to be repeated (exercise #3 and #2 two for mariner 2, and exercise #6 and #2 for mariner 3) because of technical problems with the ECPINS system during the test run. The exercises were repeated after all of the other scheduled exercises for that mariner had been completed.

## Performance Measures

### Objective measures

*Cross-track Deviation.* The Norcontrol simulation computer printed an outline chart with course lines marked for all of the ships in the exercise. It also printed out the latitude, longitude, heading, course made good and speed of the exercise ship (own ship) at ten-second intervals. These printed data were scanned into computer files and provided the movement database for each exercise.

The latitude and longitude waypoints for the start, finish, and each course change of the planned course for each exercise were provided by CMS. A computer program calculated the cross-track deviation of the actual course from the planned course at each ten-second own ship position, and then found both the average value and the rms value of this deviation over the entire exercise.

The average value for each exercise is the Cross-track deviation (CTX), and the rms value is the rms Cross-track deviation (RCTX). One data point was generated for each of these measures for each exercise and each mariner. A larger cross-track deviation means that the ship was, on the average, further away from the pre-plotted course line. A large rms cross-track deviation means that the ship wandered back and forth around the course line, rather than staying, for example, always the same distance on one or the other side of the course line.

*Closest Point of Approach.* CMS provided the latitude and longitude of channel buoys and other marks that represented hazards to navigation on each exercise. A computer program calculated the closest point of approach of the own ship to each of these marks, for each mariner and each exercise. The average of the closest point of approach to each of the landmarks in the exercise was calculated for each exercise and each mariner, and the rms value of the average closest point of approach was also calculated. The average value is the CPA, and the rms value is the RCPA. One data point was generated for each of these measures from each exercise and each mariner. A small CPA distance means that the ship "cut the corners" near the extreme edges of the safe channel, while a large CPA means the ship kept to the middle of the indicated safe channel. A large RCPA means that the ship's course varied from the middle of the channel to the edge, while a small RCPA means that the ship's course stayed more consistently near (or far) from the edges of the safe channel.

*Speed.* Speed over ground was recorded at ten-second intervals along the own ship course. The average speed (SPD) and the rms speed (RSPD) were each calculated for each exercise and each mariner. A high average speed is an indication that a navigator is confidently negotiating the hazards of a particular exercise, while a high RSPD shows that the navigator finds it necessary to make many speed changes during the exercise.

**Mariner Activity**

An experimental assistant in the simulator control room (Figure 4) monitored consoles that repeated both the radar and the ECPINS displays in the wheelhouse. He also monitored an open intercom channel to the wheelhouse and could hear all rudder and engine orders given by the mariner. The assistant recorded the time and type of each rudder and engine order, and the time and type of each mariner-controlled change in either the radar or the ECPINS display, including switching from a superimposed to a separate radar display in the choice display condition.

The sum total of all of the display console changes (displays) and all of the engine and rudder orders (orders) was recorded for each exercise and each mariner.

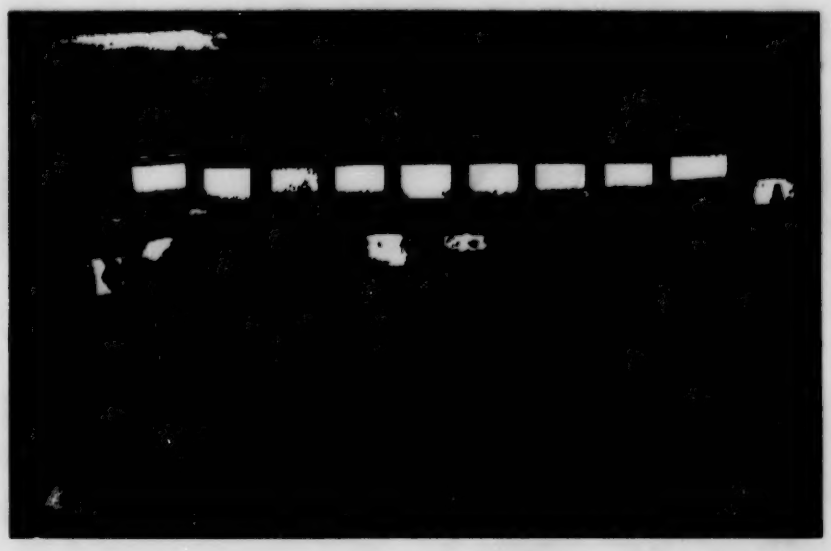


Figure 4. Simulator control room, showing the experimental assistant (left) monitoring the intercom and recording changes in the ECPINS and radar displays, and the operator (right) supervising the experiment.

A large number of display console changes means that the navigator interacted frequently with the display information source. This may be related to the information workload required by the exercise. A large number of engine and rudder commands also means that the navigator was required to interact often with the ship control system. This may also be related to the workload required by the exercise.

**Self-ratings**

After each experimental session, each mariner completed a self-rating form (Appendix C) which was loosely based on the NASA TLX workload scale (Hart & Staveland, 1988). Mariners were asked to rate each just-completed exercise for how much mental demand, physical demand and time pressure it produced. Then they were asked to evaluate their own performance, the level of overall effort required, and the degree of frustration encountered, while carrying out the exercise. Each rating was done by checking off positions on a ten-position line that was anchored at Low and High. The ratings were converted into numbers from one to ten. A single derived scale of demand was created by averaging the ratings from mental demand, physical demand, time pressure and overall effort. Performance and frustration were evaluated separately.

**Written comments**

Mariners were asked to make written comments on their self-rating forms. These were collated and are presented in Appendix D.

## RESULTS

There were few important or significant differences among the three display conditions while there were large and significant differences among the three exercise types and large and significant differences among individual mariners.

The results are summarized in terms of the two experimental manipulations of interest; display condition (choice, overlay or separate) and exercise type (RC, Rc and rC), which are presented in Table 4. Individual differences among the mariners are presented in Table 5, and the mariners' written comments are summarized in the text.

Each set of response measures (objective measures, activity and self-rating) was analyzed separately using a two-factor multivariate analysis of variance, followed by two-factor analyses of variance on each single measure where appropriate. The analysis factors were display condition, exercise type, and their interaction. Mariners were treated as a random factor, and the interactions of mariners with each experimental factor and their interaction constituted the error terms for the analyses. The analyses were carried out using routines from the SAS System for Linear Models (Littell, Freund & Spector, 1991). All statements about significance of differences between effects are based on $p < .05$ as determined by the t-test probability values for paired comparisons between the appropriate means, as generated by the SAS program.

### Display Conditions

None of the objective performance measures differed significantly across the three display types. The multivariate analysis of variance showed that the variability among the objective performance measures across display types was no greater than would have been expected by chance.

One objective measure did differ significantly across the display types, but only when the differences were evaluated within each exercise type. Figure 5a shows the relationship between average cross-track deviation (CTX), display type and exercise type. Average cross-track deviation is significantly higher over all of the Rc exercises than over the other two exercise types, but the difference among exercise types is significantly reduced under the overlay display condition. Exercise Rc cross-track deviation under the overlay condition is not significantly different from the cross-track deviation recorded under the other two exercise types (RC and rC).

Self-rated performance was significantly different across both display type and exercise type, and self-rated performance also differed significantly for each display type depending on the type of exercise. Figure 5b shows that average self-rated performance was higher in the choice display condition than the other two conditions, and also shows that under the overlay and separate display conditions, self-rated performance was significantly lower in the rC exercise type, but in the choice display condition, self-rated performance did not differ significantly across exercise types.

Multivariate analysis of the two mariner activity variables (displays and orders) gave a non-significant difference across display conditions, but display activity was significantly greater under the overlay than the choice condition when evaluated by a univariate analysis of variance.

## Exercise Types

### Objective measures

All of the objective performance measures differed significantly across the three exercise types. All of the recorded values were greater for the Rc exercises, in which heavy radar traffic was encountered in coastal waters rather than the narrow confines of a harbour or seaway. Figure 6 illustrates this effect for the objective measures. Multivariate analysis of variance of the differences across exercise conditions showed that pattern of differences across the objective measures was consistent and highly significant. A supplementary comparison evaluating the difference between the Rc condition and the average of the other two conditions: RC and rC, was also significant for all of the objective measures.

### Self-ratings

The three self-rating measures; demand, performance and frustration, all differed significantly across Exercise type (Figure 7), and a multivariate analysis of variance showed that the differences were highly significant. Demand and frustration were rated as significantly higher, and performance was rated as significantly lower, under the rC exercise conditions (little radar traffic on a complex chart), than either of the other two conditions. Performance was negatively correlated with demand ($r = -.40$, $p < .001$) and with frustration ($r = -.24$, $p < .05$), and frustration was positively correlated with demand ($r = .52$, $p < .001$), when the correlations were calculated across the 54 data points (nine exercises times six mariners) of the study.

### Activity

Neither the number of display manipulations, nor the number of rudder and engine commands, differed significantly across the exercise conditions.

P515400.PDF [Page: 26 of 67]

Table 4. Performance Measures for Display Conditions and Exercise Conditions.

| Performance measures | Display conditions | | | | | | Exercise conditions | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Choice | | Overlay | | Separate | | RC | | Rc | | RC | |
| | M | SE | M | SE | M | SE | M | SE | M | SE | M | SE |
| **Objective performance** | | | | | | | | | | | | |
| ACTX | .047 | .010 | .031 | .008 | .052 | .014 | .024 | .012 | .087 | .012 | .018 | .007 |
| RCTX | .036 | .009 | .027 | .008 | .039 | .013 | .015 | .012 | .071 | .011 | .016 | .007 |
| CPA | 4.379 | .864 | 1.259 | .733 | 8.592 | 1.246 | .349 | 1.090 | 13.697 | 1.047 | .657 | .651 |
| RCPA | 4.077 | .816 | 1.121 | .693 | 8.053 | 1.177 | .269 | 1.030 | 12.284 | .989 | .158 | .615 |
| SPD | 8.958 | .434 | 8.863 | .369 | 8.121 | .626 | 6.636 | .548 | 10.349 | .526 | 8.958 | .327 |
| RSPD | 0.950 | .215 | 1.340 | .183 | 1.449 | .310 | 1.268 | .548 | 1.348 | .261 | 1.124 | .162 |
| **Self-Ratings** | | | | | | | | | | | | |
| Demand | 4.452 | .314 | 4.577 | .281 | 4.682 | .276 | 4.247 | .314 | 3.666 | .277 | 5.799 | .280 |
| Performance | 7.838 | .393 | 6.423 | .331 | 6.583 | .345 | 7.388 | .393 | 7.298 | .347 | 6.158 | .350 |
| Frustration | 3.682 | .566 | 4.647 | .507 | 3.873 | .497 | 3.517 | .566 | 3.435 | .500 | 5.250 | .504 |
| **Activity** | | | | | | | | | | | | |
| Displays | 20.59 | 2.33 | 25.68 | 2.08 | 21.05 | 2.04 | 21.55 | 2.33 | 25.15 | 2.05 | 20.63 | 2.07 |
| Orders | 25.16 | 3.00 | 21.84 | 2.68 | 24.80 | 2.63 | 25.25 | 3.00 | 20.20 | 2.65 | 26.34 | 2.67 |

ACTX = average cross-track deviation, RCTX = rms cross-track deviation, CPA = average closest point of approach, RCPA = rms closest point of approach, SPD = average speed, RSPD = rms speed. M= mean, SE= standard error of the mean.

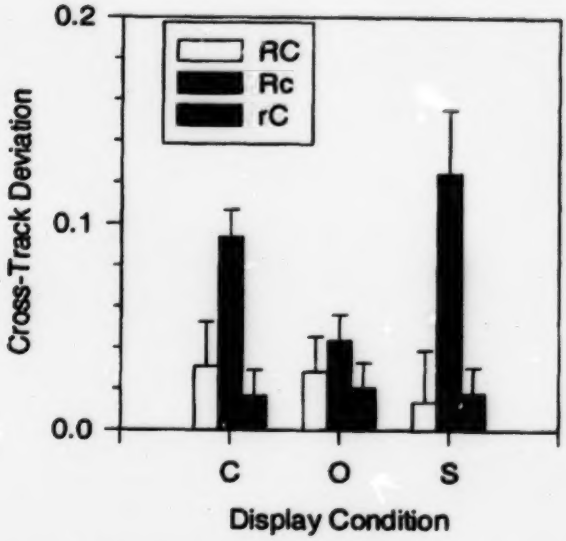



Figure 5a. Average cross-track deviation for choice, overlay and separate display conditions for three exercise types (RC, Rc and rC).

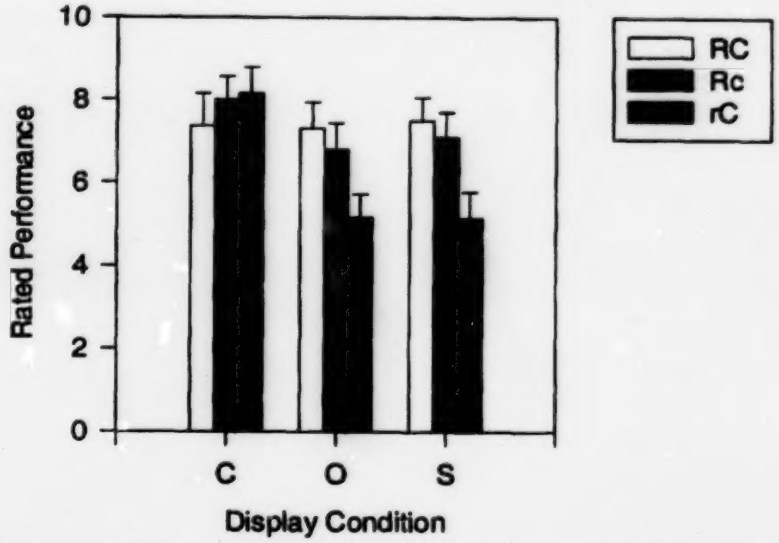



Figure 5b. Self-rated performance for choice, overlay and separate display conditions and three exercise types (RC, Rc and rC).

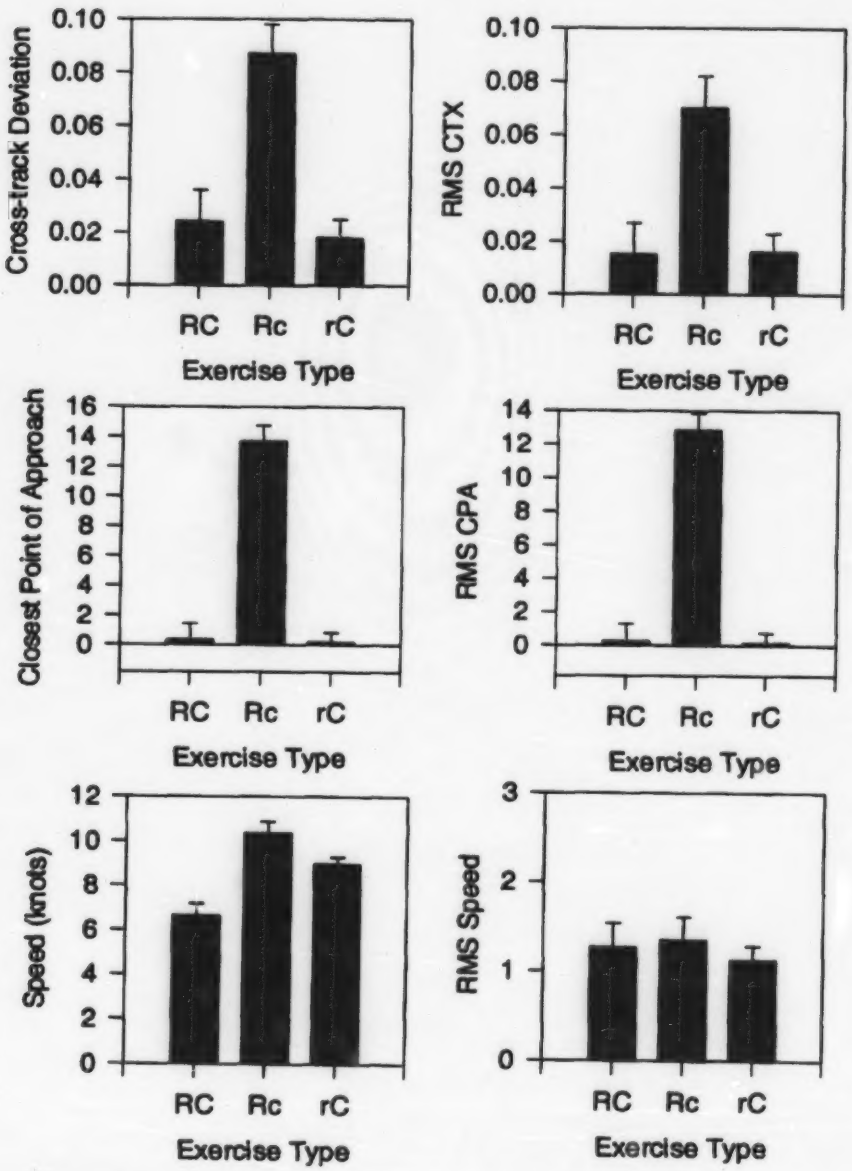



Figure 6. Objective performance measures for exercise types.

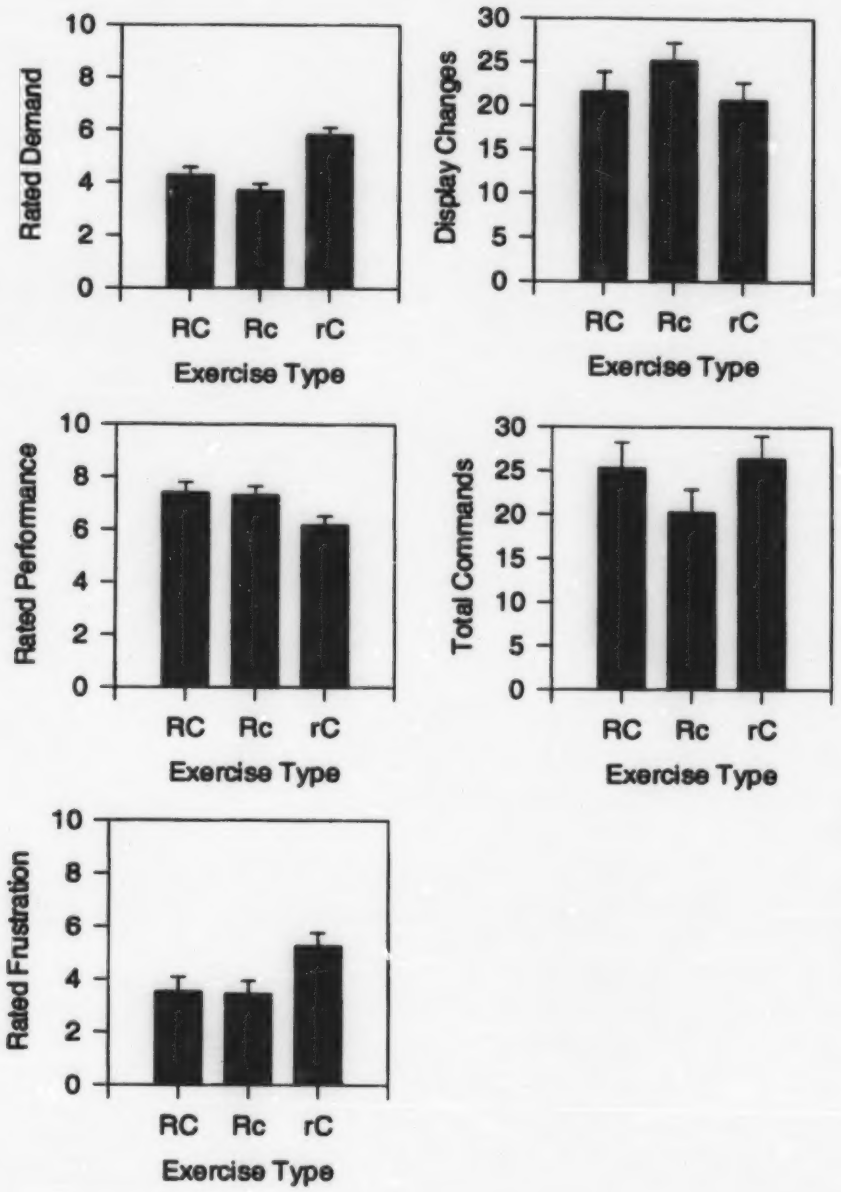



Figure7. Self-rating and activity measures for exercise types.

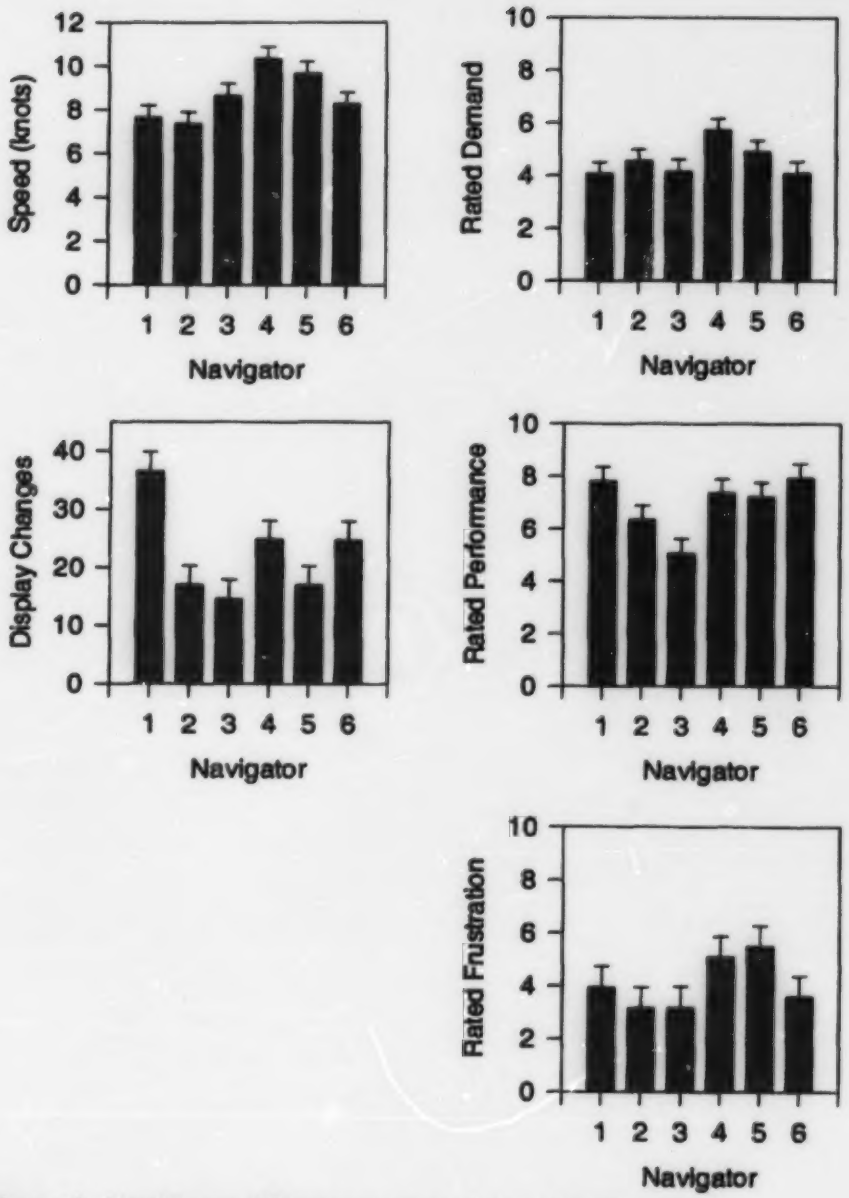



Figure 8. Significant differences in navigators' performance on objective measures, activity measures and self-ratings.

P515400.PDF [Page: 31 of 67]

Table 5. Individual Mariners' Performance.

| Performance measures | Mariners | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | |
| | M | SE | M | SE | M | SE | M | SE | M | SE | M | SE |
| **Objective performance** | | | | | | | | | | | | |
| ACTX | .047 | .011 | .041 | .011 | .039 | .012 | .030 | .011 | .017 | .011 | .059 | .011 |
| RCTX | .040 | .012 | .031 | .012 | .036 | .012 | .024 | .011 | .014 | .011 | .044 | .011 |
| CPA | 3.009 | .123 | 3.051 | .123 | 3.066 | .128 | 2.704 | .120 | 2.749 | .121 | 3.058 | .122 |
| RCPA | 2.838 | .131 | 2.837 | .130 | 2.861 | .136 | 2.439 | .127 | 2.465 | .129 | 2.885 | .130 |
| SPD | 8.260 | .457 | 7.770 | .454 | 8.413 | .474 | 10.703 | .444 | 10.448 | .451 | 8.508 | .452 |
| RSPD | 1.132 | .273 | 1.469 | .271 | 0.495 | .283 | 1.284 | .265 | 1.338 | .269 | 1.223 | .270 |
| **Self-Ratings** | | | | | | | | | | | | |
| Demand | 4.050 | .441 | 4.542 | .438 | 4.143 | .458 | 5.722 | .428 | 4.883 | .435 | 4.085 | .436 |
| Performance | 7.801 | .552 | 6.334 | .549 | 5.049 | .536 | 7.369 | .536 | 7.211 | .545 | 7.921 | .546 |
| Frustration | 3.942 | .796 | 3.150 | .791 | 3.155 | .826 | 5.093 | .773 | 5.480 | .785 | 3.583 | .787 |
| **Activity** | | | | | | | | | | | | |
| Displays | 36.539 | 3.267 | 17.029 | 3.248 | 14.514 | 3.392 | 24.876 | 3.173 | 17.003 | 3.225 | 24.690 | 3.234 |
| Orders | 26.324 | 4.214 | 23.097 | 4.189 | 19.293 | 4.374 | 21.826 | 4.093 | 23.030 | 4.159 | 30.028 | 4.170 |

ACTX = average cross-track deviation, RCTX = rms cross-track deviation, CPA = average closest point of approach, RCPA = rms closest point of approach, SPD = average speed, RSPD = rms speed. M= mean, SE= standard error of the mean.

## Mariners

### Objective measures

Independently of either display conditions or exercise conditions, some mariners drove their simulated ships faster than others, as shown in Figure 8. Mariner 4 drove his ship significantly faster than anyone except mariner 5, while mariner 5 drove his ship significantly faster than anyone except mariners 3 and 4. None of the other differences among mariners' speed were significant.

### Self-ratings

Mariner 4 rated the demand imposed by the exercise significantly higher than everyone except mariner 5. Mariner 3 rated his performance significantly lower than everyone except mariner 2. Differences among the mariners in self-rated frustration were marginally significant; only Mariner 5 rated his frustration as significantly greater than Mariner 2's (Figure 8).

### Activity

Mariner 1 made significantly more display changes than any of the other mariners, across all of the exercises (Figure 8), and mariners 4 and 6 made marginally more display changes than did mariner 3.

### Summary of written comments

Several mariners wrote that they preferred the separate radar and electronic charts, because, for example, "sailing with radar overlay through remarks and warnings printed on the chart was annoying and made it more frustrating." Another mariner said that "using e-chart without radar overlaid gives a better display, closer to using paper charts." On the other hand, a mariner commenting on a separate display condition said that "radar overlay would have reduced workload in this case as much attention had to be given to vessel being set down to port by winds."

These comments relate to the determination of the ideal level of display complexity. Too much clutter makes the display unusable; but adding radar information to a chart display, which increases display complexity, would be useful if it allows the mariner to deal with all the relevant navigation information on the same spatial display.

In the choice display condition, mariners had the option of using the radar overlay on the ECPINS display, or using the radar console, or using both. Mariners 1, 4 and 5 chose to leave the radar overlay on all the time during the three choice exercises. Mariner 1 commented, apropos of exercise 7, "the radar overlay was kept on only because it at this exercise was not in the way," but he actually kept the overlay display on continuously during all three of the choice exercises. Mariners 2 and 6 kept the overlay display off all

of the time during the three choice exercises. Mariner 3 turned the overlay on only a few seconds before the end of the exercise.

It was impossible to evaluate the dynamic display complexity as it evolved through the exercises, because there was insufficient budget available to re-run the simulations in the presence of observers whose task would be to provide a continuous real-time evaluation of the complexity of the radar and ECPINS displays. Asking the mariners to do this during their experimental runs would have prevented them from giving the necessary full attention to the piloting task.

Technical problems with the radar-chart alignment on some exercises, and resulting doubts about the chart positioning information, led some mariners to question the reliability of the electronic chart: "When comparing ECDIS and ARPA it appears that the chart has on offset of 0.1 nm almost north", and "Running over buoy on electronic chart which is seen visually and by radar 0.1 nm to starboard does not instill confidence in the system." These comments emphasize the essential reliability required of the positioning information for the chart with respect both to the radar image and the geographical world it represents.

These concerns are reinforced by a note from a CMS operator about a plotting error in exercise #2, which brought the planned ship's track too close to a pier of the McDonald bridge in Halifax Harbour (Appendix E). The point of the note was that several of the mariners navigating this exercise failed to visually observe the close approach. The bridge pier did not appear on the ECPINS chart. The concern expressed was about undue reliance on chart information to the exclusion of good watch keeping, which was particularly difficult in this exercise because of the 0.4 nm visibility.

## Discussion

Objective performance measures from the CMS study cited earlier (Mercer and Hong, 1994) show that electronic chart navigation with or without radar overlay is better than conventional paper charts and radar for navigation and piloting in confined waters and harbour approaches. Based both on the earlier CMS data and on the data from this study, there is no significant or important difference between objective performance using electronic charts with a radar overlay, and performance when the electronic chart and the radar display are presented on adjacent separate screens. If the electronic chart is reliable and complete (see the note concerning the bridge pier above), then it will allow better navigation performance than a paper chart of the same area.

The only modification of this general conclusion compelled by the evidence reported here is that for the Rc exercise type (heavy radar traffic on simple charts), the overlay display condition reduced cross-track error, which was higher under the choice and separate display conditions, to about the same level as reported in the other two exercise conditions (RC and rC).

In the earlier CMS study, mariners self-rated the navigation workload as lower, but not significantly so, in the overlay condition than in the separate condition, and both electronic chart condition workload ratings were significantly lower than the paper chart workload rating. In the present study, there was no reported difference in self-rated task demand among the three electronic chart conditions: choice, overlay or separate. However the mariners in the present study rated their own performance as significantly better under the choice display condition, despite the fact that their objective performance was not significantly better under the choice condition. They also made significantly more display changes in the overlay condition.

All of the objective performance measures as well as the self-ratings point to the Rc condition as being the least demanding. The Rc exercises (heavy radar traffic on simple charts) produced significantly larger cross-track errors, greater closest points of approach, and higher speeds than the other two exercise types, and the Rc exercises were also rated lowest (but not significantly so) on demand and frustration. Rc was effectively tied for highest rating on self-rated performance (7.30, versus 7.39 for RC). As shown by the less tightly controlled navigation performance, the mariners tended to "cut corners" relative to the plotted course when the navigation situation was objectively simpler, the perceived workload was reported as lower, and self-rated performance was higher.

The significant performance and self-rating differences found as a function of exercise type point up the increase in attention and vigilance required when maneuvering in narrow channels with fixed and transient hazards that generate complex charts (e.g. the St. Lawrence Seaway and Halifax Harbour exercises in RC and rC) as opposed to making a harbour approach or departure in more open waters with simpler charts (e.g. the Come-by-Chance and Halifax approaches exercises in Rc). Open approaches are faster, require less attention to distance from navigation hazards, and do not demand as diligent a maintenance of the planned course. In spite of the "freer spirit" of the open approaches and departures found in the Rc exercises, the overlay display condition did reduce the cross-track deviation to approximately the same level experienced in the more demanding RC and rC exercise types. I conclude from this result that the overlay condition, in which the navigator is forced to use the ECPINS with radar overlay and cannot look at a separate radar scope, coerces a more rigorous piloting discipline than would otherwise be observed when the navigator can avoid the information conveyed by the superimposition of chart and radar returns.

The rC condition (light radar traffic in a complex chart environment) showed a consistent pattern of self-rating. It was rated significantly as the most demanding and frustrating condition with the lowest self-rated performance. A possible explanation for this is that mariners feel anxious when restricted visibility (the operating condition for all of these exercises) is combined with few radar targets in a complicated navigation environment. The continued presence of radar targets in the RC and Rc conditions is a reassurance that another sensor is substituting for missing vision, and the fewer radar targets of rC may leave the mariner uncertain, frustrated and anxious about whether he will encounter targets that are not painted by the radar.

# STUDY 2: THE COMPLEXITY OF VISUAL DISPLAYS

## GENERAL INTRODUCTION

Tullis (1983) developed a measure to quantify the structural complexity (information) in an alphanumeric display. He also reviewed research that proved that measures of structural complexity were related to performance efficiency with alphanumeric displays. But neither Tullis nor anyone else had developed a graphic image-based structural complexity metric until Donderi and McFadden (1999) reported that the JPEG image compression algorithm, which is based on a discrete cosine transformation of components of an image bitmap (Salomon, 1998), generates a measure than can predict image complexity ratings of prints made from the chart, radar and overlay images used in the 1999 study. The measure was simply the JPEG compressed image file size.

Image compression algorithms are designed to reduce or to remove the redundancy from an image bitmap in order to reduce the file size, and so increase digital display and communication efficiency. The bits of a compressed file convey the minimum information available to reconstruct an image from the file. They are an indirect measure of the amount of non-redundant information in the image. The smaller the file, the less non-redundant information in the image. It is clear from the results reported in Donderi and McFadden (1999) that the quantity of non-redundant information, as measured by the compressed file size, is correlated with the subjective complexity rating of the original image. JPEG compressed image file size correlated with rated complexity for all three of the image sets used in the 1999 study (Figure 9).

One goal of the research reported here was to replicate the discovery that image complexity is related to JPEG compressed file size. The goal was met in two ways. First, the same set of images used in the previous study was re-tested for rated image complexity, and the new ratings were correlated with the previously calculated JPEG image file size. Second, two new sets of images were obtained, converted to bitmap files, and then compressed using both the standard JPEG image compression algorithm and the new JPEG 2000 standard, which uses a slightly different compression scheme. The images in these sets were also rated for complexity, and the complexity ratings compared to both the JPEG 2000 and JPEG compressed file sizes.

## EXPERIMENT 1: REPLICATION OF THE COMPLEXITY STUDY

### Method

In the previous study (Donderi, 1999) the three image types: chart, radar, and overlay, were each evaluated for complexity separately by a different group of ten observers. Each rater viewed all of the images of one type. The images were presented as mounted, high-quality ink-jet printouts of the image bitmap files, spread out together on a large table.

Research on subjective judgment has established that range and sequence effects strongly influence the quantitative outcome of a scaling study. (Poulton, 1979; Stevens, 1975). Scale anchors, that is, a small set of samples that are assigned fixed scale values, can be used to control both range and sequence effects by providing observers with "fixed points" to which they can relate the judgments of other samples. An extension of this technique called "master scaling" provides for a portable set of "fixed points" that permit comparisons to be made between different scaling studies. A different approach to reducing range and sequence effects was taken in these studies. The Proscale magnitude estimation scaling technique (Donderi, 1997) is designed among other things to reduce sequence effects, and control range effects, by allowing the observer to look at and compare all of the samples in a freely chosen sequence, while assigning scale values. This allows all of the samples to act as a collective anchor for each of the samples. Cleveland, Harris & McGill (1983) showed also that simultaneous presentation of all the samples tends to linearize the usually curvilinear psychophysical function.

The Proscale technique was used to collect the complexity rating data reported in the earlier experiment (Donderi, 1999). Although the sample groups were large (53 items), which helps to control for range effects, each observer group saw only one type of sample: either charts, or radars, or overlays, which reduces the range of complexity experienced in the sample set. Reducing the maximum possible range of complexity values across the sample will most likely increase the variability with which any number within that range will be assigned, leading to a less stable complexity judgment, and therefore one that is less likely to correlate highly with an external criterion. So in the first replication experiment, each observer saw samples from all three types of display: charts, radar, and overlays, with the goal of extending the range of complexity judgments and increasing their stability. There are other techniques to achieve this goal, but the presentation technique described here was believed to be an improvement on the previous technique.

The same image printouts were again presented to three groups of ten new observers. Each of the three image sets contained one-third chart images, one-third radar images unrelated to those chart images, and one-third overlay images unrelated to either the chart or the radar images in the same set. The three image sets were created by dividing the radar, chart and overlay images into three sets of 18, 18 and 17 images each, and then selecting the radar images from one set, the chart images from another, and the overlay images from the third set, to constitute each of the experimental sets. Thus each observer's image set contained all of the 53 topographical locations, but each location was represented by only one type of image: either a chart, or a radar, or an overlay.

In most other respects the experiment was carried out in the same way as the experiment reported in Donderi (1999). Thirty observers were recruited from among McGill undergraduate students attending summer school. Each observer was paid $ 10.00 for his or her participation. Nineteen of the observers were women. Observers were given written instructions and completed a complexity rating form for each image (Appendix F). The same ink-jet printouts were used as before, and again, all the images in a set were presented together on a large table. Each image was identified by a randomly assigned

code number. Raters were free to take as much time as needed to complete the rating form. Each observer was tested individually. The only difference was the composition of the image sets. Whereas previously each set consisted of all one type of image (radar, chart or overlay), this time each set consisted of one-third of each kind, with no overlap among the geographical regions represented. The study was carried out between May 8 and May 24, 2000. Once the data were collected, the ratings were analyzed using the Proscale multidimensional scaling and magnitude estimation scaling program (Donderi, 1997).

## Results

Figure 9 shows the relationship between the complexity ratings obtained for each type of image (chart, radar and overlay) and the JPEG compressed file size for each image. Each figure includes the complexity ratings obtained for the same image type from all three experimental groups. Thus the results plotted in each figure contain data from all thirty experimental observers, each of whom rated some of the images of that type in the context of rating other images of the two other types.

Comparison of the new graphs with the figures from the 1999 study shows that the relationship between rated image complexity and JPEG file size is as strong or stronger in the replication study as in the original.

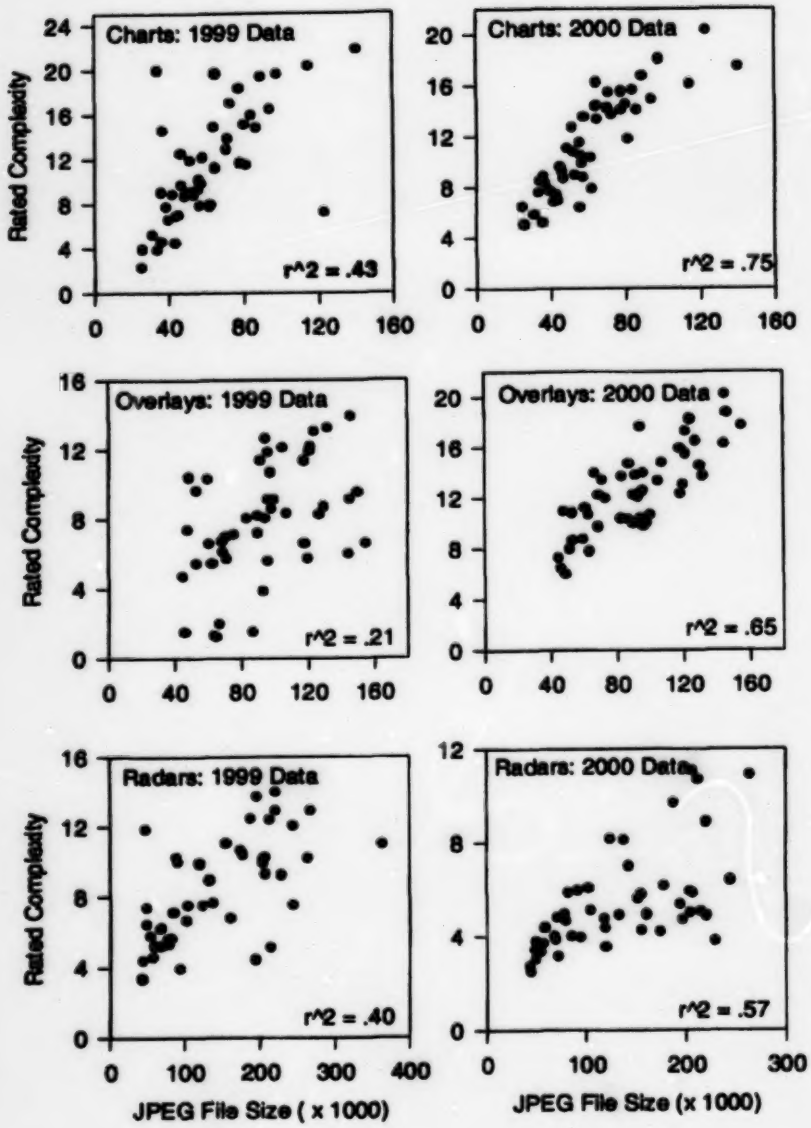



**Figure 9. Relation between JPEG compressed file size and complexity from 1999 study and from the present replication.**

## Experiment 2: Extension to new image sets

### Method

#### Images

The photography book *Earth from Above* (Arthus-Bertrand, 1999) provided a set of twenty color photographs of natural features ranging from glaciers to volcanoes. Each of these photographs was scanned into a computer and converted to a bitmap file. The bitmap file was then compressed using both the JPEG compression algorithm and the JPEG 2000 compression algorithm, a new algorithm that uses a wavelet transform instead of the discrete cosine transform of the JPEG algorithm (ISO, 1999).

Twenty web sites were captured and saved as bitmap files. These files were then also compressed using both the JPEG and the JPEG 2000 algorithms.

Figures 10 and 11 reproduce four of the nature photos and four of the web sites used in this study.

#### Observers

Twenty new McGill University undergraduates attending summer school were recruited to complete this experiment. They were paid $10.00 each for their participation. Fourteen were women.

#### Design

The bitmap file of each image in each set (Nature images and Webs) was printed out on a high resolution inkjet printer using photo-quality paper. The images were mounted on neutral gray cardboard backgrounds, and given a random identification number that appeared on the lower right-hand corner of the background. The twenty observers were divided randomly into two groups of ten. Each group rated the complexity of one of the two image sets. All the images in a set were visible together as in the previous rating experiments. Raters were free to take as much time as needed to complete the rating form. Each observer was tested individually. The study was carried out between June 28 and July 6, 2000. Once the data were collected, the ratings were analyzed using the Proscale magnitude estimation scaling program (Donderi, 1997).

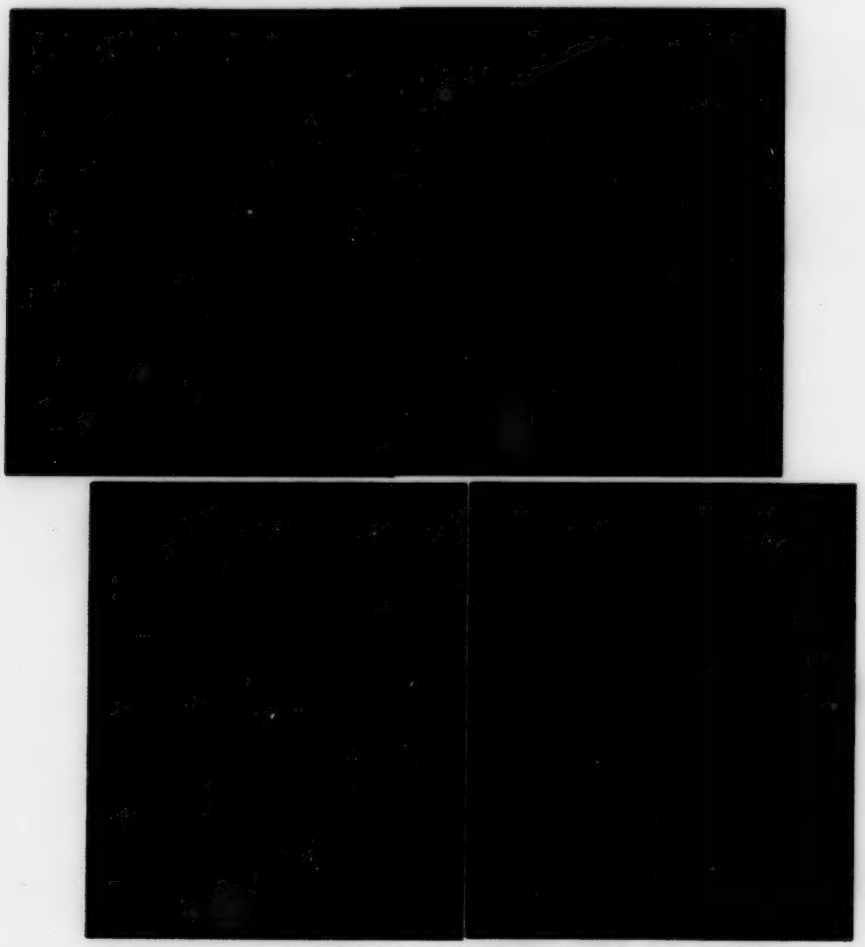


Figure 10. Four of the twenty nature photographs used in the replication study.

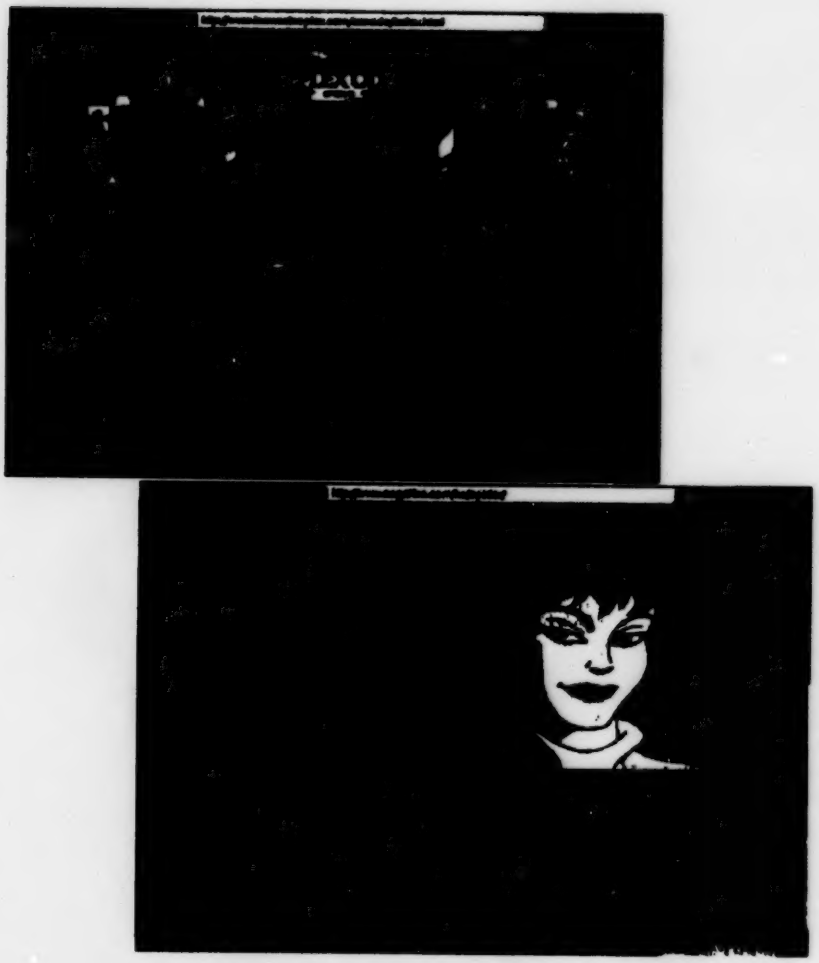


Figure 11a. Two of the twenty web sites used in the replication study.

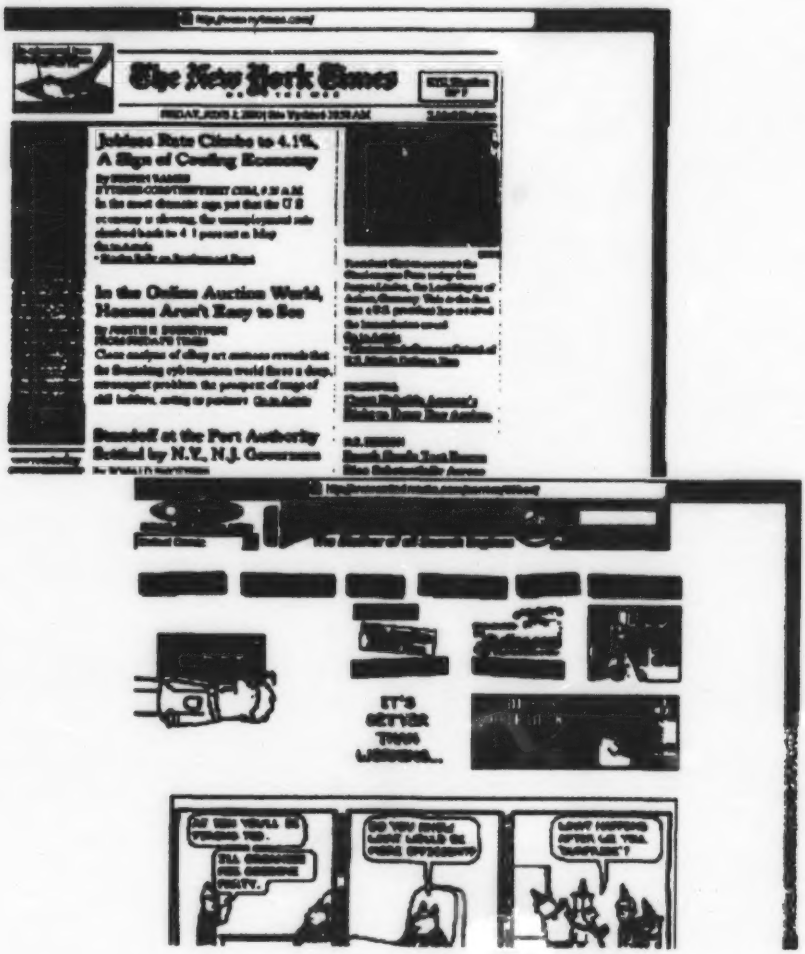

The New York Times
Jobless Rate Climbs to 4.1%, A Sign of Cooling Economy
In the Online Auction World, Houses Aren't Easy to Sell
Standoff at the Port Authority Settled by N.Y., N.J. Governors
IT'S NEVER THIS LONELY...

Figure 11b. Two of the twenty web sites used in the replication study.

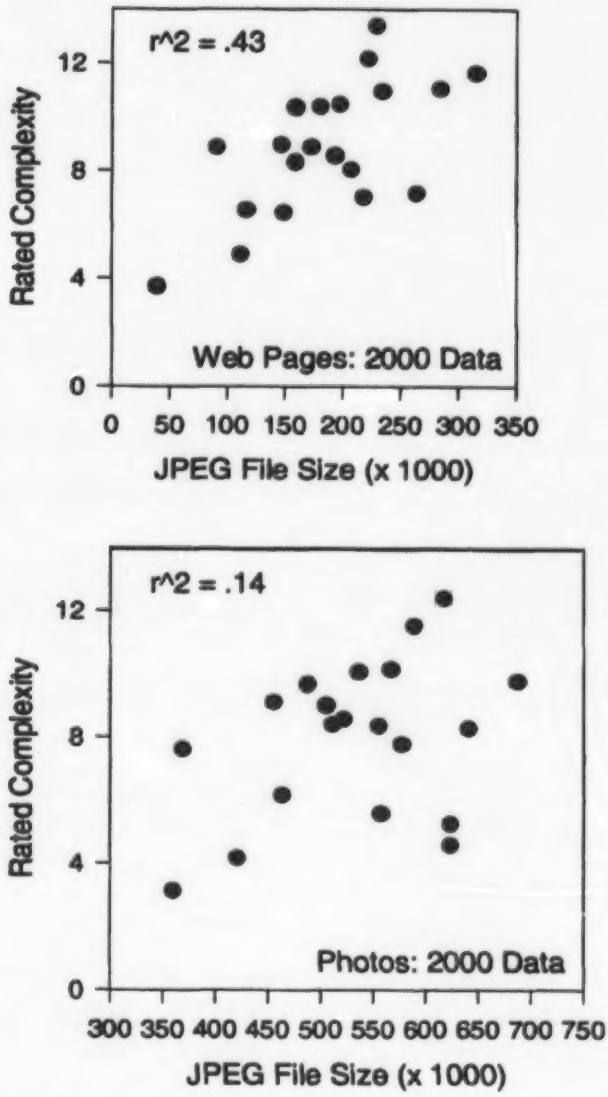



Figure 12. Relationship between JPEG compressed file size and rated complexity for the website and natural scene images.

## Results

### Compression algorithms

The new JPEG 2000 algorithm and the old JPEG algorithm produced almost identical compressed file sizes for the forty images used in this study ($r^2 = .96$). For our purposes, the two algorithms give the same result. We used the JPEG 2000 image file size results as the physical measurement for Experiment 2 of the replication study.

### JPEG file size and complexity

Figures 18 and 19 show the relationship between JPEG compressed file size and complexity for the two new image sets. There is a strong positive correlation in both sets. The results of this part of the experiment thus extend the work reported by Donderi (2000).

## Discussion

The scatter of points relating JPEG file size to complexity judgments is smaller for the chart and overlay display in the replication experiment than it was in the original experiment. This is probably because observers in the replication experiment had a wider range of images to compare. In the replication experiments all three image types were presented to the same observers, giving them a wider set of visual experiences with which to anchor their rating values. But the complexity ratings for the radar images were not less scattered in the replication experiment, so this explanation can be valid only for the items at the higher end of the complexity range: the chart and overlay images. Subjective judgments of the overlay images, which are a superposition of radar and chart images, were almost as well correlated with JPEG file size as were the judgments of chart images alone. So radar image subjective complexity, when added to the chart complexity, was adequately accommodated by JPEG file size. Human judgment of radar image complexity is not well correlated with the file size measure, perhaps because humans are better than the JPEG compression algorithm at removing the non-redundant but non-informative "noise" in a radar image[1].

There is somewhat more scatter in the relationship between JPEG file size and judged complexity for the new images: nature scenes and websites, than for the charts, overlays and radar images that were used in both the old and the new studies. There were only twenty images in each of these sets, as opposed to more than fifty in the other sets, and this relatively restricted range of complexity may have reduced the precision with which people assigned their complexity ratings. Nevertheless the relationship between JPEG file size and complexity ratings is strong and significant for these image sets. With a larger set of items of either kind (webs or nature scenes) there would be a greater range for subjective comparison and most likely a stronger relationship between JPEG file size and subjective complexity.

[1] I thank Sharon McFadden for this suggestion.

Even the web sites, which contain a great deal of text, followed the relationship demonstrated for the other more completely graphical images: the charts, overlays, radars and nature images. JPEG compression is not considered an ideal tool for the compression of alphanumeric files, but, as Donderi & McFadden (1999) showed, even for several relatively simple alphanumeric displays, the JPEG compression file size did follow the information quantities calculated by Tullis.

So far as I am aware, there is no other image-based quantitative measure that correlates with subjective complexity ratings of graphic images. The simple measure of JPEG compressed file size is correlated to subjective complexity because it is a measure of the non-redundant information content of an image. The larger the JPEG compressed file size, the greater the non-redundant information, and the greater the non-redundant information, the greater the experience of complexity. Appendix G shows that dissimilarity can be derived from complexity, and performance efficiency in the earlier static image experiments was predicted from subjective dissimilarity.

It will be important to confirm that quantitative measures of image complexity do correlate with human performance based on extracting information from images.

## GENERAL DISCUSSION AND CONCLUSIONS

The earlier CMS research summarized in this report (Mercer & Hong, 1994) showed that for navigation and piloting in confined waters, electronic chart navigation display systems are superior to navigation with paper charts. The electronic display systems produce objectively superior performance and they are experienced as requiring less work. The new research reported here shows that the choice of a radar overlay on the electronic display does not significantly influence objective performance, and also shows that mariners prefer the option of choosing to use the overlay, or not, as the situation dictates.

The comments recorded by mariners after each exercise indicate that technical reservations about the simulated system prevent mariners from giving electronic charts their full confidence until they can be assured that the ship location on the chart accurately reflects the ship's geographical location. The bridge pier problem in exercise #2 is indicative of what happens when chart information is incomplete or inaccurate.

Mariners' comments also bear on a more fundamental unresolved issue, which is the optimum amount of information on the ECPINS overlay display. The same mariners complained after one exercise that the overlay screen was too cluttered, and commented after another exercise that that the radar overlay was just right. The differences of opinions from one exercise to the other are as striking as any overall differences among the mariners.

How useful a display will be to the mariner depends on its actual information content at any moment, and that in turn depends on two things: the complexity of the chart and the radar return resulting from the geographical and the hydrographic features, and the amount of radar-reflecting traffic encountered. In this study there was radio silence and visibility was reduced to 0.4 nm, so all crucial piloting decisions had to be made well in advance of seeing other traffic, and were based entirely on radar and electronic chart information. Comments make it clear that mariners preferred the radar and electronic chart on separate displays when confronted with a cluttered single overlay display, but found the radar overlay to be quite useful when the display was not overloaded with information.

Other comments indicated that even the simple electronic chart was sometimes too cluttered to be useful. Current electronic charts have much less resolving power than paper charts. Mariners commented that reading the chart through the printed instructions was difficult, and also that the traffic separation schemes (in Halifax harbour) were unusable, presumably because they were hard to see and not because they were bad seamanship.

## Reducing Overlay Display Clutter

The ECPINS system used in this study was not the most current version of its own manufacturer's system. Some of the suggestions made here may have been met by more recent systems of this or other manufacturers. Nevertheless the issues raised apply to all electronic chart systems, and should be considered when planning any new system or evaluating any existing one.

### Separate Language Layers

Canadian electronic hydrographic charts can be de-cluttered by adding a separate layer for the French and English language instructions, which presently occupy a single display layer. The two languages are redundant, and a mariner only wants one at a time. Eliminating one of the two languages at the Mariner's choice will reduce the verbal clutter on any Canadian chart by one-half. It is a simple change that will improve the legibility of Canadian electronic charts. I recommend strongly that it be done now.

### Independent Gain Control for Land Radar Returns

Ranges and bearings of radar returns from prominent landmarks are essential when radar is being used in conventional piloting, but they are inessential when the navigator is using GPS to position the ship on an electronic chart. Radar returns from ground clutter and returns from prominent landmarks are unnecessary information when the mariner is satisfied about the accuracy of the display, and the primary purpose of radar is to provide information about ship traffic hazards and other aspects of the seaway.

An immediate reduction in the redundant and unnecessary information conveyed by a radar overlay to an electronic chart can be obtained by linking the radar overlay image with the electronic chart image so that the bitmap colors for chart areas that represent land can be used to suppress the radar overlay for those same areas. The optional ability to suppress radar returns from land areas, or to independently control their gain, will permit the navigator to eliminate this information on the overlay (but not on the radar scope) if radar is being currently used only as a guide to what is on the surface of the water, including buoys and ships. A logical exception can and should be made for returns from active radar ranges or beacons.

## FUTURE DEVELOPMENTS IN ELECTRONIC NAVIGATION DISPLAYS

Marine electronics manufacturers will add display capabilities when the market demonstrates a strong interest in them. Aside from issues of display alignment accuracy and system reliability, an increase in display resolution is probably the single technological change that can be easily sold as an improvement for marine electronic navigation displays. The next most salable improvement will be an increase in display size, because increasing both resolution and size allows the display provider to increase the total information-carrying capacity of the display without excessively increasing the information density per unit area, and so discouraging the navigator with clutter. However there are tradeoffs with larger displays; mainly the fact that less of the information can be grasped within a single visual fixation, and more fixations, and therefore more actively directed attention, will be required to process information from the same geographical area.

On the display design side, a very careful reduction of display information until a minimum basic display that has the necessary chart information presented with as little structural information (display complexity) as possible will provide the best chart foundation for a useful radar overlay. Designing this minimum electronic chart requires iterative testing and improvement until an optimum design has been achieved.

### Flat-Panel Displays

Flat-panel displays can be incorporated into efficient wheelhouse designs, whose design principles incorporating flat-panel displays were worked out very thoroughly in the nineteen-eighties. However, recently tested flat-panel displays do not have the visibility angle, the contrast range or the color gamut required to meet the IMO standard for electronic charts (Cowan, 1999). So wheelhouse flat-panel displays will not appear until the displays meet IMO standards, and can do justice to already sophisticated wheelhouse design concepts (Schuffel, 1985; Hall & Anderson, 1980; Herrman, 1977).

### Display Information Layers

Careful reduction of display information until a minimum basic display that has the necessary chart information presented with as little structural information (display complexity) as possible will provide the best chart foundation for a useful radar overlay. The use of a basic chart display with a radar overlay produces the piloting advantage of the overlay – information from both sensors on the same display – while minimizing the adverse effects described by mariners in this study -- information overload, or clutter, produced by the complexity of the chart display. Designing this minimum electronic chart and the other display layers requires iterative testing and improvement until an optimum design has been achieved. Based on the work of Burns (2000), the optimum chart display for use with a radar overlay may convey more information than heretofore imagined.

Every structural component of a display contributes to the information carried by it: for example, changes in colors, font styles, font sizes and line widths. These information-carrying components are different from the topographical information conveyed by chart contours and radar returns. They must be efficiently used so as to reduce the amount of informational noise (e.g. a type style change not associated with a change in information content), and optimize the amount of useful or usefully redundant information (e.g. a type style change associated with different types of information conveyed), in each display layer.

On the basis of the research reported here, whether to use an overlaid or a separate radar display is a choice best left to each navigator. It will depend on the current navigation situation, what else is happening, and each navigators' own preference for, and competence to use, varying amounts of simultaneously presented display information.

## Display Complexity Measurement

The second study reported here was directed at developing a quantitative metric for display complexity. Since complexity is another word for the clutter that discourages mariners from using electronic charts with or without radar overlays, a metric that measures display complexity and predicts performance on electronic chart displays will be useful in the marine environment, and in other environments that need efficient electronic displays.

The work reported here replicates the finding that JPEG compressed file size correlates with subjective ratings of display complexity. The finding was based on a retrospective analysis of data collected in a previous DCIEM-sponsored study of electronic chart navigation using static displays. This earlier work showed that the dissimilarity between the radar and associated electronic chart image was correlated with performance efficiency of electronic chart and radar images, presented both as separate

and as overlaid displays, that were studied in a question-answering format. Intermediate levels of dissimilarity led to the most efficient performance.

The radar-chart dissimilarity measure that was found to relate to efficiency in Donderi (1999) can be derived from complexity measures of each of the images: radar complexity, chart complexity and overlay complexity. This is demonstrated in Appendix G. And if we can prove that a quantitative complexity metric correlates highly with performance on these graphic displays, then we will have succeeded in extending Tulllis' (1983) metric for the complexity of alphanumeric displays to the larger universe of two-dimensional, static, graphic displays.

Since the discovery of a graphic display complexity measure is relatively recent, there has not been time to develop a complete research program designed to elicit the properties of this measure as it relates to marine navigation. I believe that the potential applications of such a measure justify carrying out the research.

## REFERENCES

Arthus-Bertrand, Yann (1999) *Earth from above*. New York: Harry N. Abrams

Burns, C. M. (2000). Putting it all together: Improving display integration in ecological displays. Human Factors, 42(2), 226-241.

Cleveland, W. W., Harris, C. S. & McGill, R. (1983). Experiments on quantitative judgments of graphs and maps. *Bell System Technical Journal*, 62, 1659-1674

Cowan, M. (1999). *Review of flat screen technologies for ECDIS displays: Plasma, LCD.* Report to the Canadain Hydrographic Service

Donderi, D. C. (2000) A practical metric for the complexity of graphic displays. *Experimental Psychology Society-BBCS Joint Meeting, Cambridge, UK, July 2000* (Abstract)

Donderi, D. C. (1997). A multidimensional scaling program for distinctiveness and similarity based on stimulus classification and information measurement. *Behavior research methods, instruments and computers*, 29(4), 549-555

Donderi, D. C. (1999) *An experimental study of marine radar and electronic chart integration.* Report to the Defence and Civil Institute of Environmental Medicine, Downsview, Ontario.

Donderi, D. C. & McFadden, S. (1999) Performance measurement of combined versus separate radar and electronic chart displays. *19th International Cartographic Conference, International Cartographic Association*, Ottawa

Habberley, J. S., Shaddick, C. A. & Taylor, D. H. (1986) A Behavioural Study of the Collision Avoidance Task in Bridge Watchkeeping. Report to the Marine Directorate, Department of Transport (UK). Warsash, UK: College of Marine Studies.

Hall, J. W. & Anderson, M. D. (1980) The U.S. Coast Guard Multi-Mission Cutter: Command, Display and Control (COMDAC). Naval Engineer's Journal, 92(35), 59-69.

Hart, S. G. & Staveland, L. E. (1988) Development of NASA-TLS (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (eds), Human mental workload. Amsterdam: North Holland.

Herrman, R. (1977). Two studies for optimizing operating bridges and their application in inland and sea-navigation. Human Factors in the Design and Operation of Ships: Proceedings of the First International Conference on

Human Factors in the Design and Operation of Ships, Gothenberg, Sweden, pps 58-68.

International Maritime Organization (undated) IMO Resolution A.819 (19): Performance Standards for Electronic Chart Display and Information Systems (ECDIS).

ISO (1999) JPEG2000 requirements and profiles version 6.0. Document # ISO/IEC JTC 1/SC 29/WG 1 N1385

Istace, H. (1977). An experimental evaluation of a "one-man control" bridge layout. Human Factors in the Design and Operation of Ships: Proceedings of the First International Conference on Human Factors in the Design and Operation of Ships, Gothenberg, Sweden, pps 167-173.

Kaufmann, R. (1990). Comparison of minimum safe and more complex displays on electronic charts. Displays, 11, 87-92.

Littrel, R. C., Freune, R. J., & Spector, P. C. (1991). SAS system for linear models. Cary, NC: SAS Institute Inc.

Mercer, R. M. & Hong, M. B. (1994) Passage Execution: A Comparison of Human Performance using Electronic and Paper Charts. Final Report to the Canadian Hydrographic Service. St. John's Nfld: Centre for Marine Simulation.

Millar, I. C. & Hansford, R. F. (1983). The 'Manav' integrated navigation system. Journal of Navigation, 36(1), 81-92.

Poulton, E. C. (1979). Models for biases in judging sensory magnitude. *Psychological Bulletin,* 86, 777-803

Salomon, D. (1998) *Data compression: the complete reference.* New York: Springer

Schuffel, H. (1985). Determining effects of bridge design on ship control. Society of Naval Architects and Marine Engineers, STAR Symposium, Norfolk, Va, pps 275-283.

Stevens, S. S. (1975). Psychophysics: Introduction to its perceptual, neural and social prospects. New York: Wiley.

Tullis, T. S. (1983). The formatting of alphanumeric displays: A review and analysis. Human Factors, 25, 657-682.

Wickens, C. D. (1992) *Engineering psychology and human performance.* New York: Harper Collins.

# APPENDIX A. DESCRIPTIONS OF EACH EXERCISE

# APPENDICES

# A, B, C, and F

# NOT FILMED

## APPENDIX B. MARINER'S BRIEFING SHEET FOR THE SIMULATION EXPERIMENT

# APPENDIX C. MARINER'S SELF-RATING SCALE

# APPENDIX D. MARINER'S WRITTEN COMMENTS

Appendix D CMS Study: comments log

| Observer | Exercise | Order | Written remarks |
|---|---|---|---|
| 1 | 6 | 2 | Sailing with radar overlay through remarks and warnings printed on the chart was annoying and made it more frustrating. Because of the display it was difficult to use a range larger than 3 miles where I would have liked to use 6 mile range on the radar, at least some of the time. |
| | 7 | 4 | The radar overlay was kept on only because it at this exercise was not "in the way". |
| | 8 | 7 | No radar on distance more than 1.5 n mile. |
| 2 | 2 | 2 | Using e-chart without radar overlaid gives a better display, closer to using paper charts. Only disadvantage I see, is neglecting to use radar or thinking e-chart is on radar overlay. |
| | 8 | 7 | In this exercise the radar overlay was off with the e-chart. Had to rely more on visuals. |
| 3 | 6 | 1 | My first experience with ecdis at speed and first experience in 17/18 years with arpa. With a programmed course of training plus the relevant experience with the units in use would feed very confident of carrying out pilotage tasks. So far, would prefer separate ARPA and ECDIS side by side. |
| | 2 | 2 | Obs 3, order 2, ex 2 . Still prefer separate radar and ECDIS. |
| | 9 | 3 | Prefer separate ARPA/ECDIS. Familiarity with equipment major. Passage planning /navigation notebook essential. |
| | 8 | 4 | Personal opinion is still leaning towards separate ARPA and ECDIS. Radar overlay tends to clutter screen. |
| | 4 | 5 | Radar overlay for short period in use appeared better than previous exercises. Screen appeared less cluttered. |
| | 3 | 6 | Should be a better way of controlling EBLs/URMs. Other than that, having separate ARPA input to ECDIS pose no problems. |
| 4 | 9 | 1 | When comparing ECDIS and ARPA it appears that the chart has an offset of 0.1 nm almost north. |
| | 1 | 3 | ECDIS somewhat unreliable due to offset at 1.5 cables to NNW. |

| Observer | Exercise | Order | Written remarks |
|---|---|---|---|
| 5 | 2 | 1 | 1. Unfamiliar with ECDIS. Being more familiar with controls would reduce workload. Didn't use many features because I didn't know what was there.<br><br>2. Radar and chart info (i.e. buoy) not agreeing causes doubt as to overall accuracy of ECDIS. Trust radar before ECDIS or wait to see visually. |
| | 7 | 2 | 1. Use of overlay allowed me to see where the other vessels were in relation to the channel. Radar alone does not provide this.<br><br>2. More knowledge of system might have me looking through menus and using other features of the system. My basic knowledge causes me to only use what I need to get the job done. |
| | 5 | 3 | 1. In reality the work level would be higher due to communications with passing vessels, traffic control, etc.<br><br>2. A concern in reality would also be the quality of GPS/DGPS info for electronic chart. Radar overlay alleviates some of this concern as long as radar & chart coincide.<br><br>3. Chart without overlay would definitely increase workload as cross checking would be more difficult (time consuming). |
| | 1 | 4 | Using radar alongside electronics chart did not have a significant effect on the workload. |
| | 4 | 5 | 1. In reality radio communications would have indicated the intentions of the vessel ahead of us and route planning could have been done accordingly.<br><br>2. Once again to have confidence in overlay, it must coincide with chart information. In such confined quarters with even less visibility, workload would increase as doubts would be raised about information presented. |
| | 9 | 6 | 1. Overlay again must match chart data. Running over buoy on electronic chart which is seen visually and by radar .1 nm to starboard does not instill confidence in the system.<br><br>2. Traffic separation scheme on the chart is unusable. |
| | 8 | 7 | Radar overlay would have reduced workload in this case as much attention had to be given to vessel being set down to port by winds. |

| Observer | Exercise | Order | Written remarks |
|---|---|---|---|
| | 6 | 8 | Unable to follow WP2 – WP3 due to vessel on stbd side at turn point. |
| | 3 | 9 | Finally I will say again that the overlay must match the chart in order to have credibility. |
| 6 | 8 | 6 | Radar overlay not working properly. Losing vessel image off chart on low ranges. |
| | 2 | 7 | Frustration level increases because of inability to zoom out on ECDIS and course line on chart brings C/L into pillar on McDonald Bridge |

# APPENDIX E. LETTER CONCERNING BRIDGE PIER AND ECPINS CHART

The exercise # 2 passage plan was designed on the simulator. The passage was then transferred to the ECPINS. The passage is inbound through Halifax Harbour.

The passage, if followed exactly, would take the mariner into or close to the support structure of the outside bridge. The structure could be detected by radar if one looked closely. It is not indicated on the ECPINS.

It was interesting to note that during at least two of the runs (observer # 1 and observer # 3) the candidate hit the structure or averted at the last moment, even though the structure was visible at 0.4 nm. It appeared that this was caused by an over reliance of the ECPINS and by not using visual confirmation. I am not sure if observers # 2, 4, 5, 6 encountered the same problem or not. Observer # 3 had to repeat this exercise because of the use of a new radar overlay version. He recollected the problem and therefore did not make the same mistake.

I pass this observation on for your information.

# APPENDIX F. COMPLEXITY RATING INSTRUCTION AND DATA FORM

# APPENDIX G. DISSIMILARITY DERIVED FROM COMPLEXITY

## THE RELATIONSHIP BETWEEN DISSIMILARITY AND COMPLEXITY (INFORMATION) IN RELATED DISPLAYS

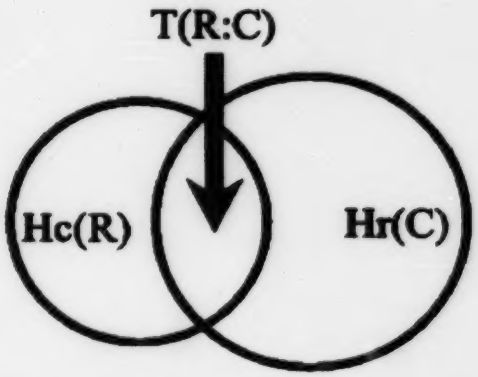
Observers in the complexity rating experiments carried out in 1999 rated four aspects of each radar-chart image pair: 1) the complexity of the radar display, 2) the complexity of the chart display, 3) the complexity of the overlay display, and 4) the dissimilarity of the radar and the chart display. Dissimilarity was related to processing efficiency for both officers and students.

Figure 1 illustrates the how complexity and dissimilarity are related in the case of a radar and a chart image. The symbol Cr refers to the complexity of the radar image. The symbol Cc refers to the complexity of the chart image. The symbol Crc refers to the complexity of the overlay image. The circles represent the overlap between complexity of the two images. The language of information theory, in which complexity is equivalent to information, is used to explain in qualitative terms how these terms are related.

The similarity of the two images is symbolized by T(R:C), which stands for the information *transmitted* between them. The dissimilarity of the two images Drc is represented by the sum of Hc(R) and Hr(C), which are the parts of the circles that do not overlap. These symbols are the information in C *independent* of R [Hr(C)], and the information in R *independent* of C [Hc(R)]. The complexity of the overlay is represented by Crc, and this symbol includes the area of both circles, because it is based on what each of the two images, radar and chart, brings to the overlay of both images. The two images bring three separate components to contribute to the complexity of the overlay: the components information in each image that is independent of the other [ Hr(C) + Hc(R)], and the third component, the information that is the same in both images; i.e. the transmitted information T(R:C).

Figure 1 represents the algebra that derives the transmitted information and the dissimilarity from the complexity (information content) of the various components. Expressed in terms of area, the area shared by both images, which is T(R:C) is found by subtracting the area including both images (the overlay complexity) from the sum of the areas of the two separate images (the complexity measures for each image). A similar algebraic manipulation gives the complement of T(R:C), which is Drc, by the following equation: $Drc = 2*Crc - Cr - Cc$.

In other words, the dissimilarity of the two images can be derived by subtracting the complexities of each single image from twice the complexity of the overlay image. The data collected from the 1999 experiments included direct estimates of all of the relevant measures. If equation (1) is correct, the direct measure of dissimilarity should be predicted from an indirect measure calculated as shown in (1). Figure 2 shows the relationship between the direct and the predicted measures of dissimilarity for the students and officers' data collected in 1999. Dissimilarity can be predicted from complexity, and objectively measured complexity (e.g. JPEG compressed file size) will prove useful in predicting performance based on image characteristics.




$$T(R:C) = Cr + Cc - Crc$$

$$Hr(C) + Hc(R) = Drc$$

Figure 1. The relationship between complexity and dissimilarity.

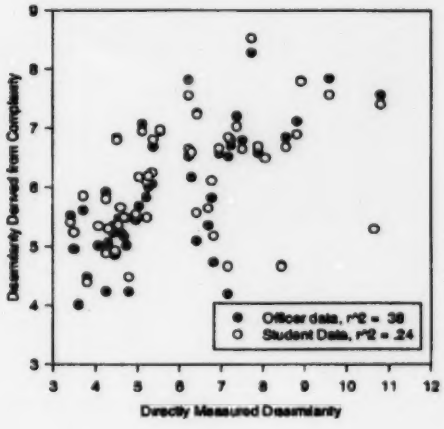



Figure 2. Relationship between directly estimated dissimilarity and dissimilarity calculated from complexity; Officers and Students' data, 1999 experiments.

| DOCUMENT CONTROL DATA SHEET | | |
|---|---|---|
| 1a PERFORMING AGENCY<br>Human Factors North Inc 118 Baldwin Street, Toronto, Ontario, M5T 1L6 | | 2. SECURITY CLASSIFICATION<br>UNCLASSIFIED |
| 1b PUBLISHING AGENCY<br>DCIEM | | |
| 3 TITLE<br>(U) The use of electronic charts and radar in marine navigation | | |
| 4 AUTHORS<br>Donderi, Don C | | |
| 5 DATE OF PUBLICATION<br>January 20 , 2001 | | 6 NO. OF PAGES<br>100 |
| 7 DESCRIPTIVE NOTES | | |
| 8 SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY<br>Sponsoring Agency<br>Monitoring Agency<br>Contracting Agency . DCIEM<br>Tasking Agency: | | |
| 9 ORIGINATORS DOCUMENT NUMBER<br>Contract Report 2001-003 | 10. CONTRACT GRANT AND/OR PROJECT NO<br>W7711-9-7561/A | 11. OTHER DOCUMENT NOS |
| 12. DOCUMENT RELEASABILITY<br>Unlimited distribution | | |
| 13 DOCUMENT ANNOUNCEMENT<br>Unlimited | | |

14 ABSTRACT

(U) Manners carried out simulated navigation exercises under three conditions: radar- electronic chart overlay, separate radar and electronic chart displays, or free choice of either the overlay or the separate display Objective performance was measured by course track, distance from navigation hazards, and vessel speed Manners also completed self-ratings of performance and difficulty There were no significant objective performance differences among the three conditions, but manners rated their own performance as higher under the free choice condition The exercises with simple chart displays were rated as producing better performance, less demand and less frustration than the exercises with complex charts A second set of experiments replicated an earlier finding that the byte size of the compressed file produced by the JPEG file compression algorithm is positively correlated with subjective judgments of image complexity for electronic chart displays and for chart-radar overlays, and extended this finding to nature scenes and web pages.

Des navigateurs ont effectué des exercices de navigation simulée dans trois conditions : avec image radar et carte électronique superposées, avec image radar et carte électronique distinctes et avec le libre choix de la superposition ou de l'affichage distinct. On a mesuré les performances objectives d'après le maintien du cap, la distance des risques pour la navigation et la vitesse du navire. Les navigateurs ont également effectué l'auto-évaluation des performances et des difficultés. Les performances objectives ne différaient pas de façon significative d'une condition à l'autre, mais les navigateurs ont jugé que leur propre rendement était supérieur dans la condition du libre choix. Ils ont indiqué que les exercices faisant appel à des cartes affichées simples permettaient d'obtenir de meilleures performances et occasionnaient moins de contraintes et de frustrations que les exercices faisant appel à des cartes complexes Un second ensemble d'expériences a confirmé la conclusion antérieure selon laquelle la taille en octets du fichier comprimé produit par l'algorithme de compression de fichier JPEG est directement proportionnelle aux jugements subjectifs quant à la complexité des cartes électroniques affichées et des affichages superposés Cette conclusion a été étendue à des scènes de la nature et à des pages Web.

15. KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) ECDIS, electronic charts; complexity; radar overlays